

**DISEÑO DE UNA METODOLOGÍA PARA LA CALIDAD DE DATOS EN LOS
SISTEMAS DE INFORMACIÓN CRM MANEJADOS EN EL CONTACT CENTER
DE EMTELCO S.A.**

MARIBEL TORRES ISAZA
MARIA CRISTINA VELÁSQUEZ BEDOYA

UNIVERSIDAD DE MEDELLÍN
ESPECIALIZACIÓN EN GERENCIA DE INFORMACIÓN
FACULTAD DE INGENIERÍAS
COHORTE 16
MEDELLÍN
2008

**DISEÑO DE UNA METODOLOGÍA PARA LA CALIDAD DE DATOS EN LOS
SISTEMAS DE INFORMACIÓN CRM MANEJADOS EN EL CONTACT CENTER
DE EMTELCO S.A.**

MARIBEL TORRES ISAZA - 43.865.474
MARÍA CRISTINA VELÁSQUEZ BEDOYA - 43.758.922

Trabajo de Grado presentado como requisito para optar al título de
Especialista de Gerencia de Información

Asesora metodología:
MARÍA DEL CARMEN SANDINO RESTREPO
Socióloga, Magíster en Sociología de la Educación

Asesor temático:
ANDRÉS OROZCO OROZCO
Administrador de Empresa

UNIVERSIDAD DE MEDELLÍN
ESPECIALIZACIÓN EN GERENCIA DE INFORMACIÓN
FACULTAD DE INGENIERÍAS
COHORTE16
MEDELLÍN
2008

CONTENIDO

	Pág.
RESUMEN	6
ABSTRACT	8
GLOSARIO	10
INTRODUCCION	14
1. REFERENTES TEORICOS	16
1.1 CALIDAD DE DATOS	16
1.2 APLICABILIDAD DE LOS SISTEMAS CRM EN UNA ORGANIZACIÓN	21
1.3 ESTRATEGIAS PARA MEJORAR LA EXACTITUD DE LOS DATOS	30
1.4 IMPACTO DE LA INFORMACIÓN EN LA ORGANIZACIÓN	32
1.5 MEJORES PRÁCTICAS FRENTE A LOS DATOS UTILIZADAS POR ORGANIZACIONES ALTAMENTE COMPETITIVAS	33
1.6 POLÍTICAS DE LA CALIDAD DE DATOS	34
1.7 ESTANDARIZACIÓN EN LA CALIDAD DE DATOS	35
1.8 DIMENSIONAMIENTO DE LA CALIDAD DE LOS DATOS	36
2. CONTEXTUALIZACION DE LA EMPRESA	37
2.1 ANTECEDENTES HISTÓRICOS	37
2.2 MISIÓN	41
2.3 VISIÓN	41
3. PROPUESTA SOBRE UNA METODOLOGIA PARA LA CALIDAD DE DATOS EN LOS SISTEMAS DE INFORMACION CRM MANEJADOS EN EL CONTACT CENTER DE EMTELCO S.A.	43
3.1 COMPONENTES	43
3.1.1 Definir un alcance	44
3.1.2 Objetivos e indicadores	59
3.1.3 Medidas de Calidad	61
3.1.4 Acciones	62
3.1.5 Recursos y Responsabilidades	70
3.1.6 Cronograma	71
3.1.7 Continuidad	71
4. RECOMENDACIONES	73
BIBLIOGRAFÍA	74

LISTA DE FIGURAS

	Pág.
Figura 1. Ciclo de vida de los datos	17
Figura 2. Premisas de CRM	22
Figura 3. ADN Emtelco S.A	39
Figura 4. Estructura Metodología	44
Figura 5. Calidad de Datos y Reglas del Negocio	46
Figura 6. Ejemplo de Modelo Entidad Relación	47
Figura 7. Reglas de Integridad de Datos	49
Figura 8. Reglas de Entidad	50
Figura.9. Reglas de Referencia	51
Figura10. Reglas de Cardinalidad	52
Figura 11. Valores Permitidos	54
Figura 12. Reglas de Herencia	56
Figura 13. Relación de Dependencia	57
Figura 14. Atributos de Dependencia.	59
Figura 15. Definición de Acciones	63
Figura 16. Ciclo de Vida de la Calidad de Datos	66
Figura 17. Chequeando la corrección de datos.	68

LISTA DE ANEXOS

Pág.

ANEXO A.	ENUNCIADO PRELIMINAR DEL ALCANCE DE LA CALIDAD DE DATOS	75
ANEXO B.	OBJETIVOS E INDICADORES DE LA CALIDAD DE DATOS	77
ANEXO C.	MEDIDAS DE CALIDAD	79
ANEXO D.	CRONOGRAMA NOMBRE PROYECTO CLIENTE	81

RESUMEN

Este trabajo de investigación se enfoca en la importancia de contar con información consistente en cualquier organización como activo fundamental para la toma de decisiones y ejecución de estrategias de cara al mantenimiento y fidelización de clientes.

Es por ello que este trabajo plantea una metodología para la aplicación del modelo de la Calidad de Datos, el cual se concentra en asegurar unos lineamientos fundamentales, las cuales facilitan alcanzar el objetivo de aplicación de la calidad de datos. La exactitud, integridad, consistencia, entereza, validez, oportunidad y la accesibilidad son las características con las cuales cualquier organización que posea un sistema de información CRM y se preocupe por tener un acercamiento personalizado y asertivo con sus clientes deberá incorporar en sus procesos internos de servicio al cliente y mercadeo aplicando una metodología como la propuesta.

Hablar de los impactos o consecuencias de una base de datos con errores es hablar de un sin número de actividades que se desarrollarían por fuera de lo planeado y que en el mejor de los casos representaría una pérdida económica por una inversión injustificada que no se pudiera recuperar, mientras que en otro caso implicaría, algo que personalmente es más impactante, la pérdida de la credibilidad de los clientes, sobre la información y el uso que se le esta dando a sus datos personales.

Partiendo de todo lo anterior, la metodología desarrollada esta conformada por siete componentes específicos, los cuales interrelacionados conllevan al cumplimiento de la calidad de datos. Inicialmente es necesario definir el alcance

para tener claridad sobre los límites del proceso y definir hasta donde deberá llegar el esfuerzo. El segundo componente es la definición de objetivos e indicadores, los cuales permitirán medir los avances del proyecto de calidad de datos, posteriormente se define el componente número tres que se enfoca en las medidas de calidad, donde se cuantifica la calidad de los datos con el fin de crear confianza entre los usuarios de la base de datos mostrando el progreso alcanzado. El siguiente componente es uno de los más relevantes para el logro de los objetivos del proyecto, es el que define las acciones específicas a seguir como lo son las auditorias, filtros, modelos de corrección y prevención. El quinto componente se enfoca en la definición de recursos y responsabilidades para medir y monitorear el proyecto, el cual se entrelaza con el sexto componente, el cronograma, el cual define el plan de trabajo que incluye las etapas y fecha de cumplimiento de las actividades asociadas al proyecto. Para finalizar, el último componente es la continuidad, donde se fundamenta el ciclo de vida de la metodología incorporando nuevos cambios que deben ser monitoreados para cuantificar el impacto de las acciones y asegurando la retroalimentación de todo el proceso de la calidad de datos.

ABSTRACT

This investigation paper is focused on the importance of counting with consistent information as a fundamental asset for the decision making process and execution of strategies for attracting new clients, maintaining them and strengthening their loyalty.

We propose a methodology for the application of the Data Quality model that is centered in assuring some fundamental parameters to facilitate the reach of the quality assurance objective. Exactness, integrity, consistency, validity, opportunity and accessibility are the characteristics which any organization that possesses a CRM information system and worries about having a personalized and effective approach with its clients must include in its internal processes of customer services and marketing, applying a methodology like the proposed.

Talking about the impacts and consequences of a data base with mistakes is talking about multiple activities that would be undertaken unlike planned and would represent, in the best case, an economic loss because of an investment that will not give any returns, or in other cases it would represent not only the economic loss, but loss of credibility from the clients, about the information and its usage.

From what has been said before, the methodology developed consists of seven specific components which are interrelated and together lead to the achievement of quality data.

First, it is necessary to define the scope to have clarity about the limitations of the process and define where the efforts should be appointed. The second component is the definition of the objectives and indicators, which will measure the progress of the quality data project. The third component is focused on the quality measures to quantify the quality of the data to create confidence among users of the data base

and showing the progress reached. The next component is one of the most relevant for reaching the objectives of the project, is the one that defines the specific actions that must be undertaken such as audits, filters, correction and prevention models. The fifth component is the definition of the resources and responsibilities to measure and monitor the project, this component links with the sixth which is the chronogram that describes the working plan and includes the stages and dates of execution of the different activities related to the project. The last component is continuity; it is the basis of the life cycle of the methodology, incorporating new changes that must be monitored to quantify the impact of the actions and assuring the feedback of the whole quality data process.

GLOSARIO

BUSINESS RULES: Es la colección de políticas y restricciones de negocio de una Organización. Las organizaciones funcionan siguiendo múltiples reglas de negocio, explícitas o tácitas, que están inmersas en procesos, aplicaciones informáticas, documentos, entre otras.

CONTACT CENTER: Es un punto de contacto que construye y conduce relaciones con los clientes y consumidores de una empresa. Este contacto se desarrolla a través del teléfono, Chat, fax, mail y sistemas de audio respuesta.

CORRECTNESS RULES: Reglas que se definen al interior de la organización para la corrección de errores.

CRM: La administración de la relación con los clientes CRM es parte de una estrategia de negocio centrada en el cliente. Una parte fundamental de su idea es, precisamente, la de recopilar la mayor cantidad de información posible sobre los clientes, para poder dar valor a la oferta. La empresa debe trabajar para conocer las necesidades de los mismos y así poder adelantar una oferta y mejorar la calidad en la atención.

Hoy existen muchos sistemas que automatizan el control de la relación con el cliente, y permiten mantener todas las herramientas comerciales a disposición de las personas que opera los sistemas para facilitar el proceso de venta.

DATA QUALITY: Calidad de datos.

DATA QUALITY RULES: Reglas que se definen para la captura de los datos en un sistema CRM.

DATAWAREHOUSE: En el contexto de la informática, un almacén de datos (del inglés *data warehouse*) es una colección de datos orientadas a un dominio, integrado, no volátil y variable en el tiempo que ayuda a la toma de decisiones de la empresa u organización. Se trata, sobre todo, de un expediente de una empresa más allá de la información transaccional y operacional, almacenada en una base de datos diseñada para favorecer análisis y la divulgación eficientes de datos.

DATOS: son los hechos que describen sucesos y entidades. Los datos son símbolos que se pueden ordenar y reordenar de forma utilizable y se les denomina información.

La importancia de los datos está en su capacidad de asociarse dentro de un contexto para convertirse en información. Por si mismos los datos no tienen capacidad de comunicar un significado y por tanto no pueden afectar el comportamiento de quien los recibe. Para ser útiles, los datos deben convertirse en información para ofrecer un significado, conocimiento, ideas o conclusiones.

FUZZY LOGIC: Se basa en lo relativo de lo observado. Este tipo de lógica toma dos valores aleatorios, pero contextualizados y referidos entre sí. Así, por ejemplo, una persona que mida 2 metros es claramente una persona alta, si previamente se ha tomado el valor de persona baja y se ha establecido en 1 metro. Ambos valores están contextualizados a personas y referidos a una medida métrica lineal.

HABEAS DATA: Es el derecho que tienen las personas de exigir confidencialidad y protección de sus datos personales almacenados en bases de datos administradas por terceros, dejando de ser un ente pasivo ante el manejo que se

le dé a su información, es algo que adquiere una enorme fuerza e importancia en la era de *Internet*.

No solamente el estatus crediticio, direcciones, teléfonos y otra información personal que se suministra en formularios cuando se hace una diligencia y que termina siendo almacenada en un repositorio en manos de una compañía o entidad estatal, deben ser objeto de la protección del recurso y derecho que ofrece la ley de ***habeas data*** al ciudadano.

Algo que la gente no conoce muy bien es que hay otro tipo de datos que se suministran sin que uno se dé cuenta. Esto sucede en *Internet* con el solo hecho de navegar por las diferentes páginas de un sitio. A medida que esto sucede, el dueño de la página va adquiriendo información sobre el navegante y si este se ha registrado anteriormente, queda perfectamente perfilado en cuanto a sus gustos y preferencias se refiere.

Por dar un ejemplo, para hacer compras en Amazon es necesario registrarse suministrando información personal. Pero esta no es tan importante como la que se recoge cuando se navega por sus múltiples páginas. Si el internauta lo que hace es solo ver DVD de ciencia ficción o libros de cocina, pues esta información queda almacenada y enlazada con sus datos personales.

INTEGRITY RULES: las reglas de Integridad se definen en el almacenamiento de los datos para que estas se cumplan y garantice que son datos correctos.

INFORMACIÓN: Información es un conjunto de datos significativos y pertinentes que describan sucesos o entidades.

MARKETING: Traducido como mercadeo, es el aprovechamiento del conocimiento a fin de aprovechar las oportunidades comparativas de la empresa a través de la comunicación con los eventuales clientes.

Es una disciplina que permite identificar, descubrir, conocer y aprovechar las necesidades de los clientes a fin de satisfacerlas de la manera más conveniente para la empresa.

ORGANIZACIÓN: Es una unidad social coordinada, consciente, compuesta por dos personas o más, que funciona con relativa constancia a efecto de alcanzar una meta o una serie de metas comunes. La organización es el acto de disponer y coordinar los recursos disponibles (materiales, humanos y financieros). Funciona mediante normas y bases de datos que han sido dispuestas para estos propósitos.

SISTEMA DE INFORMACIÓN: Los sistemas de información son procesos que se realizan sobre un conjunto de datos. Los cuales permiten recogerlos, agruparlos, analizarlos y difundirlos con el fin de realizar actividades de control y toma de decisiones.

INTRODUCCION

La mayoría de las organizaciones colombianas poseen sistemas de información para mantener las relaciones con sus clientes, sistemas CRM, pero es desde estos mismos donde se generan las principales barreras en las relaciones con estos a razón de los múltiples errores que se pueden cometer al momento de la captura de los datos, los cuales son la base para contar con la información crítica para el levantamiento de una relación altamente efectiva.

El ciclo de vida de la información en toda organización comienza con la consecución de datos que permitan entablar una relación directa con clientes, proveedores y con todos los demás actores activos que participan de la vida organizacional, por lo tanto tener datos errados, mal fragmentados, sin estandarizar o duplicados generan una conversión no fluida y con miles de problemas que pueden llegar a afectar o incluso terminar con las relaciones adquiridas después de haber invertido mucho dinero, tiempo y esfuerzo.

Es en este punto, donde los datos acarrear en sí una gran responsabilidad para transmitir correctamente un mensaje que será leído, interpretado y procesado, con el cual se llevará a cabo la toma de decisiones y se generará nueva información, la cual finalmente se convertirá en un ciclo que creará impacto a la organización; lo que lleva a concluir que los datos incorrectos, incompletos, insuficientes, irrelevantes o simplemente con mala calidad, crearán un ciclo deficiente y con iguales impactos.

Es en este escenario, donde se fundamenta la necesidad de contar con una metodología clara de calidad de datos en los sistemas de CRM's, ya que se ha identificado que por la baja calidad en los datos las organizaciones incurren en costos operativos superiores a los reales por las continuas inconsistencias y reprocesos al momento de implementar diferentes estrategias enfocadas a sus clientes y/o servicios.

Partiendo de lo anterior, este trabajo de investigación se enfoca en diseñar una metodología para garantizar la calidad de datos en los sistemas de información CRM manejados en un *Contact Center*, el caso de estudio específico que se empleara será Emtelco S.A.

Emtelco S.A es una empresa que ofrece servicios de *Contact Center* a nivel nacional, la cual opera en la ciudad de Medellín y cuenta con grandes clientes como UNE EPM Telecomunicaciones, EPM y Colombia Móvil, entre otros. A todos ellos les ofrece un portafolio de servicios que incluye toda la administración de la relación con los clientes de los servicios ofrecidos por dichas empresas. Es por esto que se ve la necesidad de diseñar esta metodología con la cual se pretende contar con unos lineamientos claros que permitan asegurar la calidad de los datos y posteriormente la interpretación y análisis de estos como insumo principal de todas las áreas de las empresas cliente de Emtelco S.A.

1. REFERENTES TEORICOS

A continuación se describen los temas fundamentales que se serán trabajados para apoyar el diseño de la metodología a desarrollar:

1.1 CALIDAD DE DATOS

La calidad de datos es uno de los fundamentos del éxito en las organizaciones. El tener acceso a información exacta y completa es fundamental para la toma de decisiones estratégicas. Los datos con calidad apoyan y fortalecen virtualmente todas las funciones de negocios y son especialmente importantes para las iniciativas que involucran:

- Servicios al cliente, comunicaciones y manejo de relaciones.
- Mercadeo y ventas.
- Manejo de riesgos.
- Planeación estratégica.

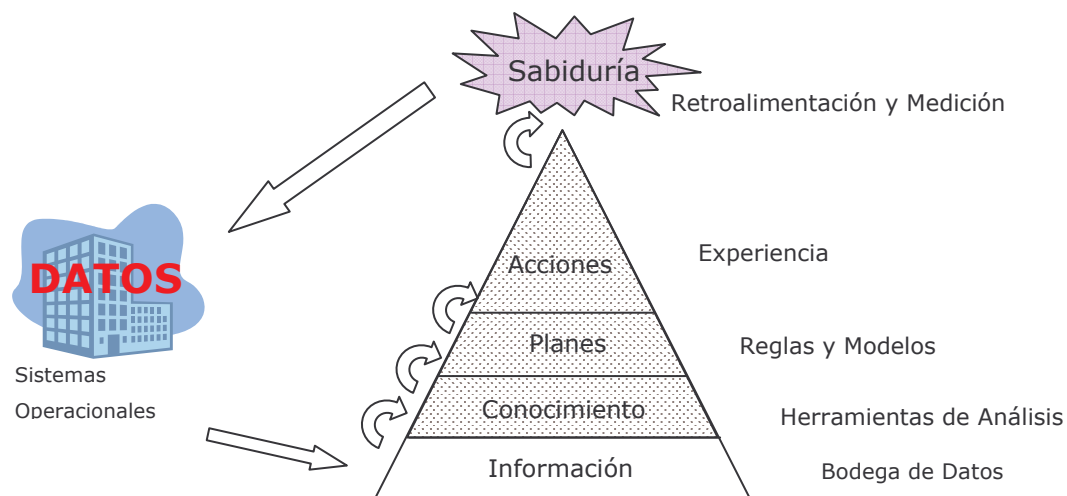
Los datos, están llevando en sí una gran responsabilidad para transmitir correctamente un mensaje que será leído, interpretado y fundamental para la toma de decisiones que impactará sobre otros que a su vez generarán nuevos datos, para generar nueva información que finalmente se convertirá en un ciclo de impacto global; lo que permite concluir que los datos incorrectos, incompletos, insuficientes, irrelevantes o simplemente con mala calidad, llevarán a crear un ciclo deficiente y con iguales impactos.

En el mundo actual, en la era de la información, se puede afirmar que la información es la base fundamental de la economía empresarial, y sobre esta expresión podríamos basar la siguiente analogía:

“Si la información es la moneda de la nueva economía, los datos son una materia prima crítica y necesaria para tener éxito, como una refinería toma el petróleo crudo y lo transforma en diferentes nuevos productos; las organizaciones usan los datos para generar multiplicidad de activos de información. Estos activos forman la base de los planes estratégicos y acciones que determinan el éxito organizacional.”¹

Los datos se convierten en información en una bodega de datos. El conocimiento se adquiere por medio de las herramientas de análisis identificando modelos en la información que serán usados para desarrollar planes de negocios. Las compañías ganan sabiduría por la retroalimentación de sus campañas y repitiendo el ciclo.

Figura 1. Ciclo de vida de los datos



Fuente: Presentación Calidad de Datos Orbitel S.A. E.S.P. Dirección de Gestión de la Información. Andrés Orozco O. Mariluz Tejada Z. 2006

¹ Presentación Calidad de Datos Orbitel S.A. E.S.P. Dirección de Gestión de la Información. Andrés Orozco O. Mariluz Tejada Z. 2006.

¿Qué es lo que puede estar mal?

- **Falta de rutinas de validación:** muchos errores en la captura de datos pueden ser prevenidos usando rutinas de validación de la información que ingresa por la Web y en los sistemas operativos.
- **Válidos, pero no correctos:** algunas rutinas de validación no logran capturar errores cuando el dato capturado es válido, pero es inválido.
- **Desigualdad en sintaxis, formatos y estructuras:** integrar diferentes fuentes de información implica que todos los formatos y estructuras estén iguales y asegurar una correcta captura.
- **Cambios inesperados en los sistemas fuentes:** un cambio en el sistema original de la información puede provocar cambios en los sistemas secundarios y reportes.
- **Redes de interfaces:** los cambios en las complejas redes de interfaces y los cambios que estas pueden sufrir.
- **Falta de comprobaciones de Integridad:** no todos los sistemas verifican la información que están cargando desde otros sistemas de información.
- **Sistemas mal diseñados:** poner en marcha sistemas operativos en tiempos muy cortos, puede provocar malos diseños u omisiones de gran impacto para la calidad de los datos.
- **Error de conversión:** un cambio en el programa de migración de datos o en la interfase puede ocasionar errores en cientos de miles de registros.

Las características necesarias y mínimas para hablar de Calidad de Datos son:

- **Exactitud:** Debe representar la realidad.
- **Integridad:** Se debe contar con toda la información crítica para el negocio
- **Consistencia:** Los datos deben ser los mismos en todos los sistemas de información.
- **Entereza:** Debe tener el nivel de detalle requerido para el tipo de decisión al cual está destinado.
- **Validez:** Debe servir verdaderamente a las personas a la cuales se les está proporcionando.
- **Oportunidad:** Debe estar disponible en el momento que se necesite.
- **Accesibilidad:** Fácil consulta.

La mala Calidad de los datos y sus impactos: Hablar de los impactos o consecuencias de una base de datos con errores es hablar de un sin número de actividades que se desarrollarían por fuera de lo planeado y que en el mejor de los casos representaría una pérdida económica por una inversión injustificada que no se pudiera recuperar, mientras que en otro caso implicaría, algo que personalmente es más impactante, la pérdida de la credibilidad de los clientes, sobre la información y el uso que se le esta dando a sus datos personales.

Para hablar y visualizar un poco mejor estos impactos, se analizarán las consecuencias que se podrían presentar partiendo de errores, que a diario se podrían llegar a presentar:

Ejemplo: En un punto de contacto se comete un error, un error pequeño, sencillo, que se podría tomar como mínimo. Este ejemplo evidencia como un error tomado como insignificante conlleva hasta la pérdida de un cliente:

- Se digitó mal el nombre del cliente
- Cuando el cliente vuelve a contactarse con el *Contact Center*, lo llamo por el nombre que no es.
- En la bodega de datos se guarda el nombre del cliente errado, quedó María y el cliente se llama Mario.
- Se saca el informe del número de mujeres y hombres que se tienen en la base de datos.
- Se crea una estrategia para:
 - Enviarles un día gratis para un salón de belleza y tienen gratis Cepillado de Pelo y Maquillaje. El premio sale impreso en una carta y dice al final: Es personal e intransferible.
 - Enviarles un correo electrónico agradeciendo por el uso de los servicios.
 - Se hace un convenio con una empresa de maquillaje y se les envía a todas las “mujeres” de la base de datos un *kit* completo de belleza con una carta personalizada.
- El cliente recibe toda esta mala información.
- Por la falta de respeto y la ofensa, que considera el cliente por parte de la empresa, no vuelve a realizar consumos para la empresa.

Lo anterior muestra, como un error calificado como insignificante en la etapa de captura de los datos, se puede llegar a convertir en una causal para la pérdida del cliente. Como este caso, hay muchos otros que tienen impactos mayores donde la organización puede afectar su buen nombre y su rentabilidad.

Otros casos comunes en la operación de un *Contact Center* son los siguientes:

- **Mala digitación del correo del cliente:** aorozco@jotmail.com.co; aorozcoo@hotmail.com.
- **Suscripción de un plan equivocado:** PLAN PEM CALI IND 2002 y el correcto era: PLAN PEM BOGOTA IND 2002.
- **Género equivocado:** era “Femenino” y se colocó “Masculino”.
- **Subtipo de caso que no era:** era “Resarcimiento Cliente Alto Valor” y se colocó “Resarcimiento”-
- **Mala digitación de la dirección:** era “CI 45D 75 27” y se digitó: “CI 45 65 27”
- **El cargo del contacto mal seleccionado:** era “Gerente” y quedó “Mensajero”.
- **Digitación incorrecta de la fecha de cumpleaños:** era “Agosto” y seleccionó “Abril”.
- **Error en el número telefónico:** era “043251783” y se colocó “043251873”
- **Omisión de información:** se deja el campo vacío.

1.2 APLICABILIDAD DE LOS SISTEMAS CRM EN UNA ORGANIZACIÓN

Las organizaciones alrededor del mundo comparten una dedicación al Manejo de las Relaciones con los Clientes (CRM). Miles de millones de dólares han sido empleados en sofisticados sistemas CRM. Sin embargo, la realidad es que la mayoría de las iniciativas CRM han producido resultados desfavorables o han sido abandonados por completo.

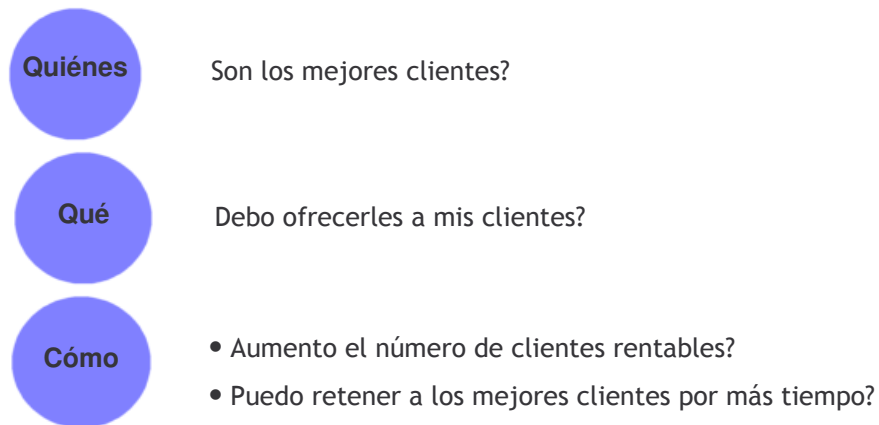
La razón principal es que mientras muchas organizaciones han enfocado su presupuesto y su atención significativamente en las iniciativas CRM, han descuidado los sistemas que apoyan al CRM, las bases de datos centradas en clientes: Archivos de Información de Clientes (CIFs), Sistemas de Información de Clientes (CISs) y Almacenes Operacionales de Datos (ODSs), que alimentan al sistema CRM con los datos críticos de los clientes.

Los datos en su base de datos de clientes son la sangre vital de su sistema CRM. Sus iniciativas CRM pueden ser sólo tan efectivas y confiables como la exactitud y plenitud de los datos de clientes que están disponibles para estas iniciativas.

El CRM y la Calidad de Datos (*Customer Relationship Management*)

Es la ciencia de identificar, atraer y retener los mejores clientes, para generar un crecimiento rentable y sostenible tanto para el cliente como para la organización. Para ello se debe tener en cuenta las siguientes premisas:

Figura 2. Premisas de CRM



Fuente: Presentación Calidad de Datos Orbitel S.A. E.S.P. Dirección de Gestión de la Información.
Andrés Orozco O. Mariluz Tejada Z. 2006.

Una estrategia de CRM podría crear una visión única del cliente que permita:

- Integrar la información del cliente de múltiples puntos de contacto.
- Generar estrategias de segmentación.
- Eliminar producción innecesaria y los gastos de envíos, por registros duplicados o información insuficiente.
- Identificar grupos familiares y otras relaciones, por medio de los correctos agrupamientos de los datos.
- Corrección y actualización de datos en línea, por medio de los sitios *web* y demás sistemas de interacción con el cliente.
- Incrementar el valor de los clientes, enfocando los esfuerzos en los más rentables.²

Estas, y muchas otras estrategias de CRM no serán exitosas, si no se cuenta con una base de datos confiable y depurada.

A continuación se detallan algunas estrategias enfocadas a la calidad de los datos:

- Establecimiento de una “**Matriz de Definiciones**”, la cual explique para cada uno de los datos dentro del sistema, todos los parámetros necesarios que lo pudieran definir y así entregar a cualquier persona un concepto completo de la razón de ser y el uso que se le da a cada uno de los datos dentro de la organización, los procesos y los diferentes sistemas de información.
- Reportes predefinidos, que permitieran **monitorear el crecimiento de la base de datos**, mostrando para cada uno de los datos una estadística básica de los datos ingresados, cuales eran correctos y cuales incorrectos; basados en unas

² Presentación Calidad de Datos Orbitel S.A. E.S.P. Dirección de Gestión de la Información. Andrés Orozco O. Mariluz Tejada Z. 2006.

definiciones básica. Otro de los informes genera una base de datos de los **clientes no actualizados** en un determinado tiempo o que no han tenido ninguna transacción en su información.

- Un **proceso** implementado y documentado, donde existe un “**Dueño de los Datos**” y es la persona responsable de autorizar cualquier modificación sobre el sistema operativo, con el fin de controlar los cambios que se le hace a los datos de los clientes y que finalmente se verán reflejados en los sistemas, las bases de datos y la información de la compañía.

Existen otras actividades de apoyo para la Calidad de la Base de datos; algunas de ellas son:

- Línea de Telemarketing Calidad.
- Capacitación y concursos para los Puntos de Contacto.
- Devolución de correspondencia.
- Actualizaciones sobre la base de datos.
- Material Publicitario.
- Optimización de la base de datos para Migración de sistemas.
- Desarrollo de productos.

Línea de Telemarketing Calidad:

Campañas de salida desde el *Contact Center*, donde la única finalidad es la actualización de los datos personales de los clientes.

Las bases de datos que se deben trabajar son las recibidas como devolución de correspondencia, errores en el sistema que implicaran tener un contacto con el cliente, por errores internos donde el cliente se vea impactado, generando así una

llamada proactiva informándole al cliente sobre lo sucedido, el resarcimiento ofrecido y se actualizaban sus datos.

Devolución de correspondencia:

A raíz de los envíos masivos de correspondencia: Cartas de Bienvenida, Cartas de Modificación de servicios, Facturación directa, Material Publicitario; se puede pactar con la empresa proveedora del servicio de envío, la creación de una base de datos con la información de las devoluciones generadas en cada envío.

La información de estas devoluciones, se puede confrontar con la información del sistema. Los resultados de esta validación se dividen en datos correctos con la información del sistema y datos incorrectos que evidencian errores, los cuales se envían a la línea de Telemarketing Calidad para ser actualizados.

Aspectos a resaltar: La devolución deja de ser una simple pérdida de dinero y se convierte en un insumo importante para la Calidad de la base de datos.

El indicador de devoluciones se convierte en un indicador de Calidad de la base de datos utilizada para el envío.

Material publicitario:

Una de las estrategias de fidelización de clientes es el envío de material publicitario, en el cual se solicite el diligenciamiento de los datos personales de los clientes y su posterior devolución a la empresa para recibir algún beneficio previamente definido de acuerdo a la estrategia puntual.

Este material se utilizará para la actualización de los datos de los clientes en la base de datos.

Aspectos a resaltar: Una nueva actualización de datos sin tener contacto con el cliente y aprovechando el material de un evento de mercadeo para optimizar la Calidad de la base de datos.

Se obtienen datos importantes como la dirección y teléfonos actualizados por el cliente.

Desarrollo de productos:

En el proceso de creación de nuevos productos, deberá existir una etapa de prefactibilidad, donde se invite a todas las áreas y se les exponga el producto y los procesos que éste impactará. En esta reunión se hacen los aportes necesarios evitando que los nuevos procedimientos atenten sobre el sistema CRM y sobre los datos que allí se capturan. Lo anterior evita errores por desconocimiento a raíz de soluciones complejas o desarrollos sobre el sistema, que se pueden evitar al presentar la mejor alternativa para ejecutar el mismo proceso sin impactar el sistema CRM.

Capacitación y concursos para los puntos de contacto:

En los diferentes puntos de contacto del *Contact Center*, se tiene un índice de rotación que implica que las capacitaciones tengan que ser continuas. Es por esto que constantemente se debe monitorear la necesidad de renovar las capacitaciones en temas de Calidad de Datos a todos los asesores.

Estas capacitaciones mostrarán desde el punto de vista teórico la definición de la Calidad de datos y llegarán finalmente a presentar los impactos y costos que la mala Calidad le implica a la compañía.

Igualmente se pueden diseñar concursos en los cuales los asesores participen con el objetivo final de incentivar la mayor y mejor actualización de los datos. Los concursos mueven un interés especial y generan un tiempo atípico en el cual se ve un crecimiento acelerado de la base de datos y del interés por el tema. Se debe saber manejar la información que se entrega en este tiempo para evitar que se quede en un concurso que pasa y no en una cultura que se debe conservar.

Actualizaciones sobre la base de datos:

En muchos casos, los errores de la base de datos no se pueden corregir desde el sistema CRM, pues las reglas de validación no permiten algunas modificaciones; lo cual implica un cambio directamente sobre las bases de datos y las tablas que soportan el sistema CRM.

Aunque estos procedimientos no son recomendables, muchas veces son la única alternativa de solución. Para esto se realiza una descripción escrita de todos los cambios que se deben realizar, el motivo y los impactos que dicho cambio implicaría. Este formato se envía a todas las personas de servicio al cliente que se relacionen con el sistema CRM para que el cambio sea aprobado y finalmente aplicado.

Para estos cambios es importante mostrar los beneficios esperados, pues muchas veces el cambio puede ser grande, pero el beneficio puede ser aún mayor

Optimización de la base de datos para migración de sistemas:

Para la migración de sistemas, es importante, revisar la base de datos que se va a migrar antes de ejecutar cualquier procedimiento. En estas revisiones se tienen en cuenta todas las variables necesarias para el nuevo sistema, las reglas de validación que tiene el nuevo sistema y las que tiene el sistema anterior.

Aspectos a resaltar: En estas revisiones es muy probable encontrar casos que no se hubieran detectado antes debidos a que en el día a día no se suele trabajar con todos los registros de la base de datos, sino puntualmente con registros que cumplen con determinadas condiciones.

En un momento de migración de sistemas, se podría analizar la posibilidad de depurar la base de datos y eliminar los registros incorrectos, no migrándolos al nuevo sistema. Para esto es necesario documentar bien cuales serían los datos que podrían quedar por fuera, el criterio de eliminación y los impactos de no tener estos datos en el nuevo sistema.

Controles básicos sobre el sistema operativo:

Se deberán crear unas reglas básicas de Calidad de datos, las cuales se deben cumplir por una norma jurídica nacional, proceso interno, definición de mercadeo o definición del negocio, y dichas reglas deberán verse reflejadas en el sistema CRM y en la base de datos final.

Los sistemas CRM permitirán programar controles especiales sobre algunos campos que realice la validación en el momento en el cual se están ingresando los datos, asegurando que desde el momento cero, los datos están siendo capturados de acuerdo a las reglas previamente establecidas.

Aspectos a resaltar: Un control sobre un campo puede ser muy útil para una regla de negocio, pero a su vez puede entorpecer otra unidad de negocio, por lo tanto antes de ser desarrollados deben ser validados con todas las reglas vigentes.

Si el dato que está siendo controlado se integra con otros sistemas de información, se deben revisar los impactos que se pudieran presentar.

Seguimiento a la calidad de datos en el sistema CRM

Los indicadores son mecanismos que permiten medir o realizar un seguimiento a los datos y a la información obtenida de estos. A continuación se muestran algunos indicadores que pueden ser aplicados:

- Variación de la base de datos de clientes.
 - Variación de Campos diligenciados correctamente.
 - Registros corregidos.
 - Cuadre de cifras entre los sistemas.
-
- **Variación de la base de datos de clientes:** Se debe extraer del sistema CRM el comportamiento de la base de datos en un tiempo determinado, con el fin de analizar alguna variación ya sea en aumento o en disminución de registros de información de los clientes.

Lo anterior con el fin de conocer el total de la base de datos que se maneja en la compañía y el comportamiento en un período de tiempo determinado, conociendo así el impacto de los diferentes eventos sobre la base total de clientes.

- **Variación de campos diligenciados efectivamente:** Este indicador mide sobre el total de registros de clientes en la base de datos, el total de clientes que quedaron con determinado dato diligenciado correctamente y por ende cuantos quedaron con el campo nulo o mal diligenciado.

Esto permite determinar la cantidad de campos faltantes o errados que están en la base de datos, detectando los errores más frecuentes y corrigiéndolos reactivamente a corto plazo.

- **Registros corregidos:** Se medirán constantemente los campos corregidos sobre los campos errados, hasta llegar a un 100% de corrección, previniendo así que queden datos errados en el sistema y el crecimiento del error en el tiempo que dure el desarrollo de la corrección del error.

Lo anterior permite medir la efectividad de las estrategias desarrolladas para la corrección de los campos detectados con errores.

- **Cuadre de cifras entre los sistemas:** Este indicador debe ser cercano al 0%, ya que está midiendo la diferencia entre los sistemas de información. Esta verificación se hace diariamente y por diferentes dimensiones, pero se tomará como medida principal el cuadre de cifras por franja horaria, ya que es una de las dimensiones más detallada.

El objetivo de este indicador es asegurar la consistencia de los datos en los sistemas de información de la compañía.

1.3 ESTRATEGIAS PARA MEJORAR LA EXACTITUD DE LOS DATOS

Existen diversos mecanismos para garantizar la exactitud en los datos manejados en los sistemas CRM de una Compañía, a continuación se detalla el método *Schober*, el cual ilustra sobre algunos aspectos a tener en cuenta para la calidad de los datos:

- **Auditoria de Calidad de Datos.** Se debe realizar un diagnóstico de los problemas de calidad de datos desde el principio. Implantar cualquier iniciativa de gestión basada en datos sin entenderlos primero, es como tratar de curar sin la comprensión de los problemas. Con servicios de Auditoría normalización de bases de datos de Schober, se examina la estructura de la calidad de datos para encontrar si la información si corresponde con las necesidades requeridas.

Los servicios de auditoría y calidad de datos de Schober automáticamente identifican los aspectos de la **calidad de datos** sobre los que se desea actuar:

- Estadísticas básicas, frecuencias y rangos.
- Identifica múltiples formas de representar el mismo contenido.
- Descubre y valida formatos y errores repetitivos.
- Identifica y valida datos redundantes y duplicados.
- Define los parámetros de configuración necesarios para los siguientes procesos de calidad de datos.

➤ **Escenarios de Aplicación**

- Calidad de Bases de Datos: asegurar la calidad postal de cada registro.
- Data Entry: ayuda online/batch, muestra de alternativas.
- Actualización periódica de BD: normalización de bases de datos; cambios en vías, CP's, población.

➤ **Deduplicación de Registros:** Los servicios de Deduplicación de Schober incluyen una funcionalidad sofisticada basada en tecnología *fuzzy logic*, tolerante a errores, la cual permite entender la importancia de elementos diferentes dentro de las bases de datos de la empresa. De esta forma es posible fusionar la información identificada dentro de o a través de diferentes fuentes de datos. Schober proporciona una variedad de servicios de deduplicación, incluyendo:

- Deduplicación/localización sobre nombre/razón social, apellido y dirección postal, y campos adicionales, basado en parámetros de experiencia de Schober.
- Exactitud: es posible de encontrar duplicados de baja similitud.
- Integración y consolidación de información existente en diferentes bases de datos y/o sistemas de información.

Escenarios de aplicación:

- Fusión de registros de bases de datos: CRM / *DataWarehouse*.
- Eliminación de registros duplicados: Duplicidad interna.
- Preparación Campañas de *Márketing*: Múltiples bases de datos.
- Búsqueda Rápida: *Contact Center*, Información escasa, 118AB.
- Chequeo Automático de Duplicados: *DataEntry*, ERP's.
- Cluster de Direcciones: Agrupación con base en criterios comunes y normalización de bases de datos.

1.4 IMPACTO DE LA INFORMACIÓN EN LA ORGANIZACIÓN

En todo el ciclo productivo de cualquier compañía, la información juega un papel fundamental que puede llegar a ser positivo o negativo de acuerdo al manejo que se le de a la misma. A continuación se describen los principales impactos que se pueden llegar a presentar:

- Campañas de mercado fallidas: Una campaña direccionada a un nicho específico de clientes y debido a la mala calidad en los datos en este criterio específico, causa que los esfuerzos de la campaña sean desviados y por lo tanto no se obtengan los resultados esperados.
- Pérdida de clientes: a raíz de una mala experiencia, un cliente puede decidir cambiar de proveedor o de empresa que le presta algún servicio determinado, ya que se siente desatendido.
- Pérdidas económicas: costos de correspondencia, de envío, de transporte, de impresión, de compra de enseres y productos específicos, los cuales están justificados para campañas diseñadas con información incorrecta.

- Reprocesos: una mala captura de datos en el centro de contacto, genera múltiples reprocesos al interior de toda la cadena productiva de la compañía, ya que un error puede generar un procesamiento de datos inconsistentes, por ejemplo en el proceso de facturación.

1.5 MEJORES PRÁCTICAS FRENTE A LOS DATOS UTILIZADAS POR ORGANIZACIONES ALTAMENTE COMPETITIVAS

Tomando como referencia las grandes compañías colombianas, se tomarán los casos de éxito de Orbitel (ahora EPM Telecomunicaciones) y Protección. Todas ellas invirtieron mucho tiempo y dinero en el tema de la calidad de los datos, pues le apostaron a este tema para mejorar las relaciones con sus clientes obteniendo beneficios directos en su rentabilidad.

Orbitel identificó graves falencias en la calidad de los datos existentes en sus CRM's y se preocupó por mejorar dicha calidad, ya que estaba incurriendo en errores, sobrecostos operativos y en afectación de su imagen a raíz de los resultados obtenidos en sus campañas de telemarketing, las cuales a pesar de su buena planeación y enfoque no impactaban al nicho de clientes adecuado y no se obtenían los resultados esperados.

A partir de lo anterior, decidieron cambiar e implementar nuevos procesos enfocados todos ellos a mejorar y garantizar la calidad en los datos. Cada contacto con el cliente se convertía en una oportunidad de actualización y de captura de datos.

Asimismo, Protección, una empresa líder en el país en Pensiones y Cesantías ha sabido utilizar muy bien la información del cliente y generando campañas de CRM muy bien direccionadas y potencializando al máximo toda la información que tienen. Por ejemplo, en el año 2007 diseñaron una campaña basada en el

conocimiento adquirido de sus clientes, en la cual utilizando la figura de un cuento felicitaba a sus socios y en dicho cuento retomaban todos estos datos para personalizar un relato de acuerdo a los hábitos, estudios, empresa donde trabaja, estado civil, fecha de cumpleaños, etc. Esta estrategia, aparte de fidelizar a sus clientes, permitió que muchos clientes se motivaran a actualizar sus datos.

1.6 POLÍTICAS DE LA CALIDAD DE DATOS

Antes de hablar de políticas de calidad de datos, cada empresa debe tener muy claras las políticas de privacidad de los datos, pues es el primer factor que se le debe asegurar a los clientes, de hecho, cuando se ingresan los datos en alguna página *web* o en alguna empresa se puede observar en alguna parte del sitio un aviso que indica al usuario que los datos ingresados en dicho sitio son regidos por unas políticas de privacidad. Estas políticas pueden indicar que los datos son de uso exclusivo para la compañía o incluso que la compañía es la dueña de los datos y se los podrá vender a terceros. En la parte legislativa existe el "*Habeas data*" que es una política que regula a las empresas de servicios para la conservación y correcto uso de los datos de los clientes.

Por otro lado, existen las "Políticas de Calidad de Datos"; las cuales se definen al interior de la compañía y se basan más en las condiciones o criterios procedimentales que se les da a las personas que ingresan los datos (asesores) o que incluso, en muchos casos y mucho mejor, son "Políticas" que se configuran directamente en los sistemas CRM, de tal forma que en el momento en que se infringe una política el sistema la puede corregir o puede generar un mensaje de error al usuario

1.7 ESTANDARIZACIÓN EN LA CALIDAD DE DATOS

La persona, el equipo o el área responsable de la calidad de los datos están directamente relacionados con áreas como Mercadeo, Servicio al Cliente, Reclamos, Facturación e Informática, entre otros. Todas las áreas con procesos de cara al cliente tendrán requerimientos sobre los datos, se darán cuenta de algunos errores o solicitarán algunas mejoras. Se debe tener en cuenta que cada área, cada proceso y cada persona de la compañía se preocupan por un grupo diferente de datos. Es decir, el que manda cartas se preocupa por la dirección y el nombre del cliente, el que manda correos electrónicos se preocupa es por el *e-mail* y el nombre del cliente y al de contabilidad no le importa para nada estos datos, sino que le interesa saber si el cliente tiene o no tiene una factura o un cobro. En este caso, dos personas, encargadas de dos procesos diferentes se preocupan por el nombre del cliente y cada uno puede tener requerimientos de calidad de datos sobre el mismo campo en el sistema de información; uno querrá que todo se grave en mayúsculas y nombres y apellidos juntos, y el otro lo querrá separado y sólo con la primer letra en mayúscula.

De igual forma en áreas de tecnología o relacionadas con procesos técnicos (Informática o Facturación) encontrarán en sus sistemas errores que les afecten sus procesos o quizá tengan la necesidad de hacer el cambio de uno de sus sistemas (Migración de plataformas), lo cual implica un trabajo grande de validación de la calidad de los datos, lo que se convierte en un requerimiento muy grande, ya que las validaciones deben ser más exhaustivas y se vuelven dependientes del sistema origen y del sistema destino.

En conclusión. El responsable por la calidad de los datos será el centralizador de estos requerimientos y así sea el área técnica (Informática) quien los tenga que ejecutar, sólo los debería hacer si el responsable de calidad de datos lo aprueba. Lo más importante en este tema es la centralización y evitar que todas las

personas de la compañía tengan la libre opción de pedir un cambio sobre los datos; pues esto puede estar afectando a otro proceso.

1.8 DIMENSIONAMIENTO DE LA CALIDAD DE LOS DATOS

El dimensionamiento inicial y básico, es el análisis de los campos principales y actuales de la base de datos; los cuales se deben estar gestionando y para los cuales ya existirán algunos procesos de depuración ya establecidos y ya dimensionados a partir de la historia (los trabajos que se hayan hecho sobre estos campos). Un segundo dimensionamiento surge a partir de las necesidades de las áreas y sus planes de trabajo, pues si se conoce que un área realizará durante el año algún trabajo especial con un campo que no es principal o que será un campo que se adicionará a la base de datos, se debe generar un análisis de los criterios que se deberán revisar y los procesos que se deberán ejecutar para asegurar su calidad. Todos estos planes de calidad se dimensionan bajo modelos estadísticos que permiten determinar según la volatilidad de los datos, el volumen, la representatividad y la importancia, cual debe ser la periodicidad de revisión o cual debe ser la muestra adecuada para la revisión.

2. CONTEXTUALIZACION DE LA EMPRESA

2.1 ANTECEDENTES HISTÓRICOS

Emtelco S.A. es una empresa prestadora de servicios de *Contact Center*, que permite generar relaciones de valor con los clientes, con un servicio cálido y un importante servicio de soporte tecnológico que permite desarrollar productos a la medida de las necesidades de cada cliente: servicios básicos de entrada, salida, IVR y marcador predictivo, que permiten atención a clientes, televenta, cobranza, investigación, encuestas, entre otros.

El negocio se ha venido desarrollando exitosamente, ofreciendo un servicio con un recurso humano amable, eficiente, integrado a la tecnología que ofrece ventajas que generan confianza para una compañía.

Actualmente la empresa cuenta con 12 clientes: Almacenes Éxito, Bristol, Colombia Móvil, Confiar, Edatel, EPM, EPM Bogotá, Gilat Colombia GNC, MAPFRE, Municipio de Medellín, UNE, Periódico el Mundo y Ultraviajes, atendiendo de esta manera empresas tanto del sector público como del privado. Emtelco está establecido en la ciudad de Medellín con cinco sedes operativas y una administrativa.

Las políticas de calidad de la compañía están enfocadas al compromiso con nuestros clientes es atender su negocio y satisfacer sus necesidades con la fortaleza de una cultura organizacional con base en una óptima calidad de los servicios brindados, mediante procesos eficientes, apoyados en la tecnología y la experiencia del recurso humano.

El objetivo de EMTELCO S.A *Contact Center* es ser la unidad de servicios compartidos estratégica de EPM Telecomunicaciones, que sirva de canal de atención no personalizada para sus usuarios finales de productos y servicios y de las demás empresas del Grupo Empresarial EPM y los clientes de base instalada, con altos estándares de calidad, eficiencia operativa y rentabilidad; apoyando desde la estrategia de *Contact Center* los siguiente procesos:

- Atención al cliente No Personalizada
- Telemercadeo
- Retención
- Cobranza

La propuesta de valor de EMTELCO S.A *Contact Center* consiste en posicionarse como el canal de contacto principal entre el Grupo Empresarial EPM y sus clientes, al integrar y desarrollar esquemas de servicios innovadores facilitando el flujo de la información, prestando asesoría en el momento oportuno, ágil y eficientemente, con lo cual se fortalecerá el esquema de relacionamiento y fidelización con los clientes actuales y potenciales de las empresas del Grupo EPM y sus filiales.³

En la actualidad las líneas del Grupo Empresarial EPM se concentran en los servicios de atención al cliente, por lo tanto la propuesta de valor se desarrollará en torno al perfeccionamiento de estándares de alta calidad en las líneas de atención al cliente y el desarrollo e implementación de servicios automatizados y de telemercadeo, cobranzas y retención, que permitan rentabilizar los procesos de los empresas del Grupo Empresarial EPM.

³ Presentación Planteamiento estratégico Emtelco S.A. Dirección General. 2007

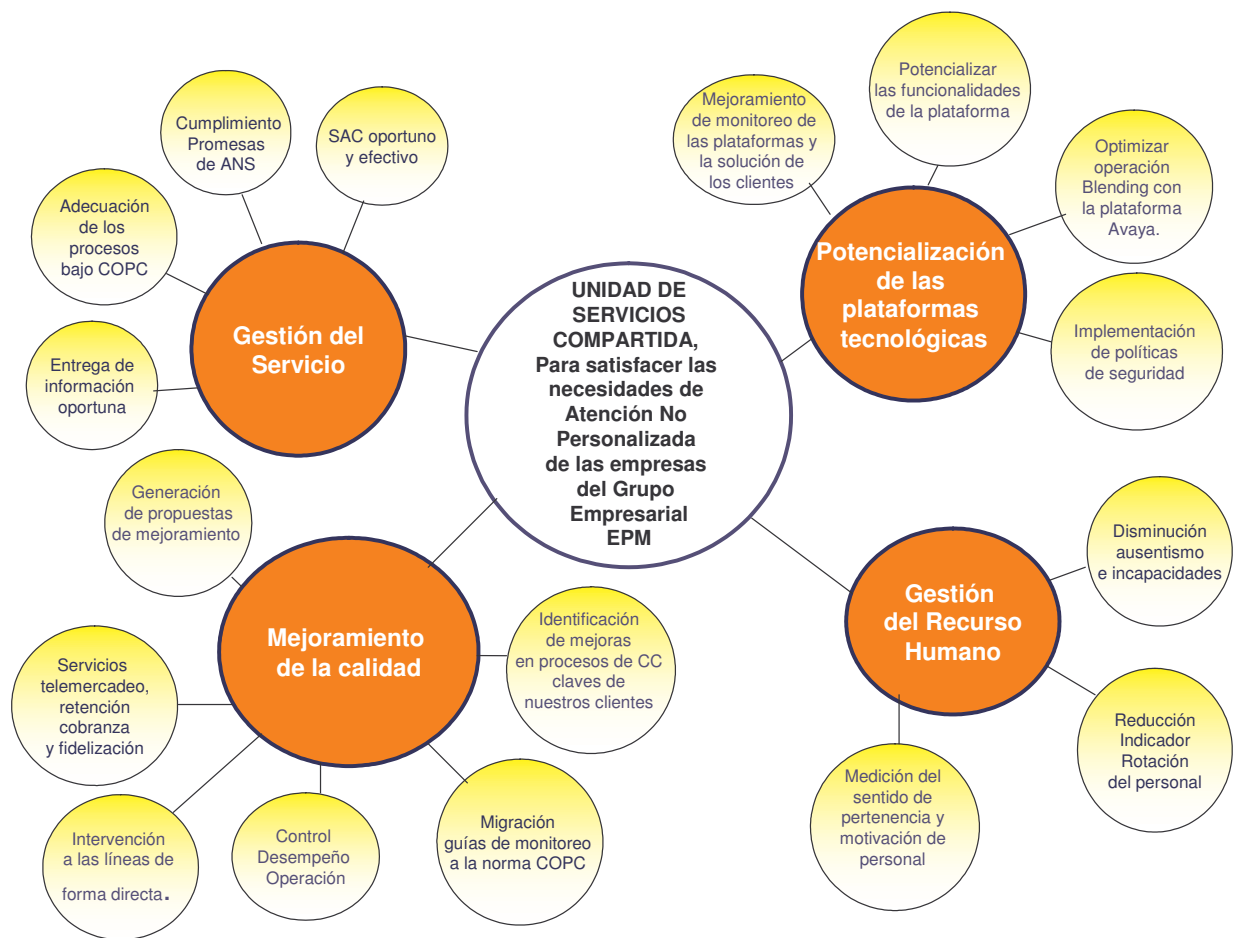
Emtelco escindido decide adoptar los valores corporativos establecidos para la matriz EPM Telecomunicaciones.

Los valores son los siguientes:

- **Enfoque en el cliente:** vivimos por la confianza que nuestros clientes ponen en nosotros y trabajamos para comprender, anticipar y responder de manera memorable a sus necesidades.
- **Enfoque en resultados:** tenemos objetivos retadores y claros, nos gusta medirlos y cumplirlos a tiempo.
- **Innovación:** creemos en nuestras iniciativas y somos creativos para resolver problemas efectivamente, satisfacer mejor las necesidades de nuestros clientes y hacer la vida más simple y placentera.
- **Agilidad:** actuamos más rápido para ajustarnos a las necesidades de los clientes y del mercado.
- **Trabajo en equipo:** Tenemos objetivos comunes y todos aportamos respetuosa y armónicamente para el logro de los mismos.
- **Responsabilidad social:** somos conscientes de nuestra responsabilidad por el mejoramiento social y económico de las comunidades en las que actuamos y procuramos generar el mayor impacto para su sostenibilidad y progreso.

En la siguiente figura se detalla el ADN para soportar esta propuesta de valor actual de Emtelco:

Figura 3. ADN Emtelco S.A



Fuente: Presentación Planteamiento estratégico Emtelco S.A. Dirección General. 2007.

EMTELCO S.A. tiene la infraestructura necesaria para desarrollar los servicios que el mercado de los *Contact Center* ofrece, para lo cual cuenta con dos plataformas de *Contact Center* que permiten atender operaciones robustas como UNE y líneas de tamaño mediano como EPMBogota.

Para las líneas de tamaño mediano, es decir, líneas con un máximo de 120 puestos, se potenciará la plataforma *Interactive Intelligence*. El mercado objetivo de esta plataforma son las líneas de tamaño medio, ya que atender operaciones de tamaño superior, por su arquitectura de múltiples servidores, que se encargan de atender porciones de los agentes, fraccionaria las líneas,

conllevar a ineficiencias en la operación desde el punto de vista de la ocupación y niveles de servicio. Sin embargo, para operaciones medianas Interactive Intelligence ha demostrado estabilidad, un mejor desempeño y una mayor personalización.

Las líneas robustas se atenderían a través de la plataforma Avaya, la cual es líder en el mercado y ha comprobado su estabilidad en operaciones a nivel mundial. El estudio de *Frost & Sullivan* 2004, concluyó que en el mercado americano de *contact centers* Avaya, es el líder por su estrategia de comunicaciones IP, reconocimiento de voz, desarrollos IVR y calidad el monitoreo.

Adicionalmente, la arquitectura de Avaya esta diseñada para tener un único núcleo central desde donde se administran todas las interacciones (llamadas entrantes, llamadas salientes, *blending*, *chat*, *email* y *fax*), extensiones y asesores, sin necesidad de fraccionar la operación, convirtiéndolas por ende en operaciones eficientes desde el punto de vista tecnológico.

2.2 MISIÓN

Trabajar con cultura de servicio basada en el conocimiento del cliente y el respeto por los compromisos establecidos, generando valor y garantizando la satisfacción de sus necesidades.

2.3 VISIÓN

En el año 2008, superar las expectativas y necesidades de los clientes garantizando el uso eficaz y eficiente de nuestros recursos, a través del desarrollo permanente de nuevas y mejores formas de hacer nuestro trabajo. ⁴

⁴ Presentación Planteamiento estratégico Emtelco S.A. Dirección General. 2007

3. PROPUESTA SOBRE UNA METODOLOGIA PARA LA CALIDAD DE DATOS EN LOS SISTEMAS DE INFORMACION CRM MANEJADOS EN EL CONTACT CENTER DE EMTELCO S.A.

Retomando de la definición de calidad de datos vista en el capítulo 1 en los referentes teóricos, la calidad de datos es uno de los fundamentos del éxito en las organizaciones y tener acceso a información exacta y completa es fundamental para la toma de decisiones estratégicas, a partir de esto se plantea la siguiente metodología que apunta a asegurar la calidad de datos en los sistemas CRM manejados en el *Contact Center*.

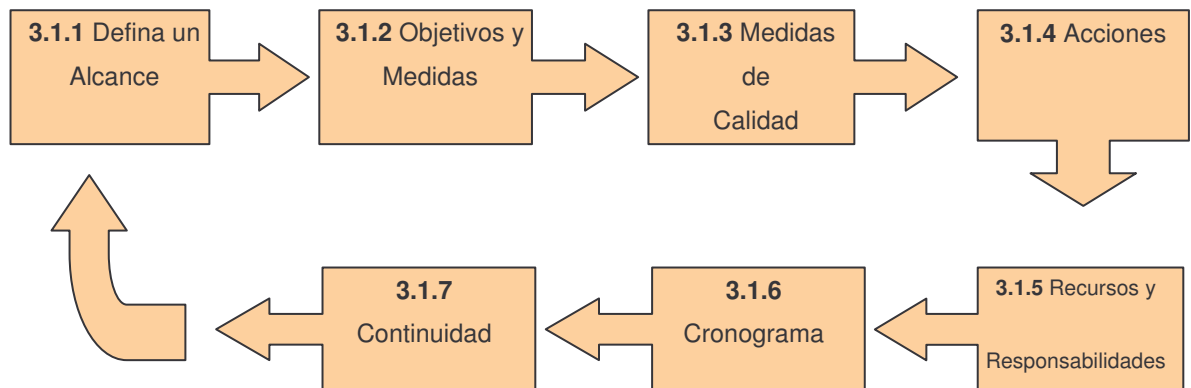
Para efectos de estudio, se planteara inicialmente el marco de referencia a seguir para definir cada uno de los pasos que conformaran la metodología y posteriormente se detallaran estos pasos con sus respectivas actividades, procesos, subprocesos, roles, recursos, acciones, reglas y formatos.

Para comenzar, se debe tener presente que calidad de datos no es necesariamente corregir errores. Los errores en los datos es sólo una parte de la ecuación de calidad de datos y minimamente se debe garantizar la exactitud, la integridad, consistencia, entereza, validez, oportunidad y accesibilidad de los datos. Teniendo clara esta premisa, a continuación se diseña la metodología.

3.1 COMPONENTES

A continuación se describen los componentes que conformarán la estructura de la metodología a plantear:

Figura 4. Estructura Metodología



Fuente: *DATA CLEANSING: Delivering high quality warehouse data.* J. Thomann

3.1.1 Definir un alcance: Es necesario conocer a profundidad los límites del proceso y definir hasta dónde debe llegar el esfuerzo a realizar. Para ello se debe dar respuesta a las siguientes preguntas:

- ¿Cuáles son las reglas del negocio?
- ¿Cuál es la fuente de los datos?
- ¿Cuál es la estructura de la base de datos?
- ¿Cuál es el sistema CRM a impactar?
- ¿Cuáles son los datos a evaluar?
- ¿Cuáles son las reglas de integridad?

Las respuestas a estas preguntas serán el punto de partida para la definición del alcance e indicarán los lineamientos sobre los cuales girará toda la metodología a seguir. Esta introducción del negocio, permitirá tener claridad en cómo se administra la información de los clientes, que CRM utilizan y las especificaciones de los datos importantes y de mayor impacto en su negocio, para definirlos como prioritarios para realizarle calidad a dichos datos.

Toda esta información será recopilada a través de una serie de formatos preestablecidos que facilitarán el posterior análisis para la definición final del alcance. Estos formatos serán incluidos en este trabajo como anexos al documento final.

El anexo A. “Enunciado Preliminar del Alcance de la Calidad de Datos” permitirá consignar toda la información requerida para tener un conocimiento general aplicable a la metodología.

A continuación se describe cada uno de los campos que solicita el formato anteriormente mencionado:

- **Descripción general del negocio:** En este campo se debe realizar una descripción general que incluya apartes de la misión, visión, objetivos estratégicos y sector en el cual se desempeña la empresa en análisis.
- **Reglas del negocio:** En este campo se deben relacionar todas las reglas de calidad de datos y de limpieza y transformación de los mismos.

En la primera se debe especificar las reglas de corrección que se enfocan en el contenido de los datos, es decir, el contenido de los datos debe contar con todas las características necesarias para proporcionar una vista confiable del negocio según las necesidades de los usuarios funcionales de dichos datos.

Asimismo, se deben incluir las reglas de integridad de los datos que se centran en la estructura de los mismos. La integridad de los datos se concentra en la estructura de los datos para que cumplan todas las características requeridas para proporcionar una vista confiable del negocio según lo requerido por los usuarios finales de los datos.

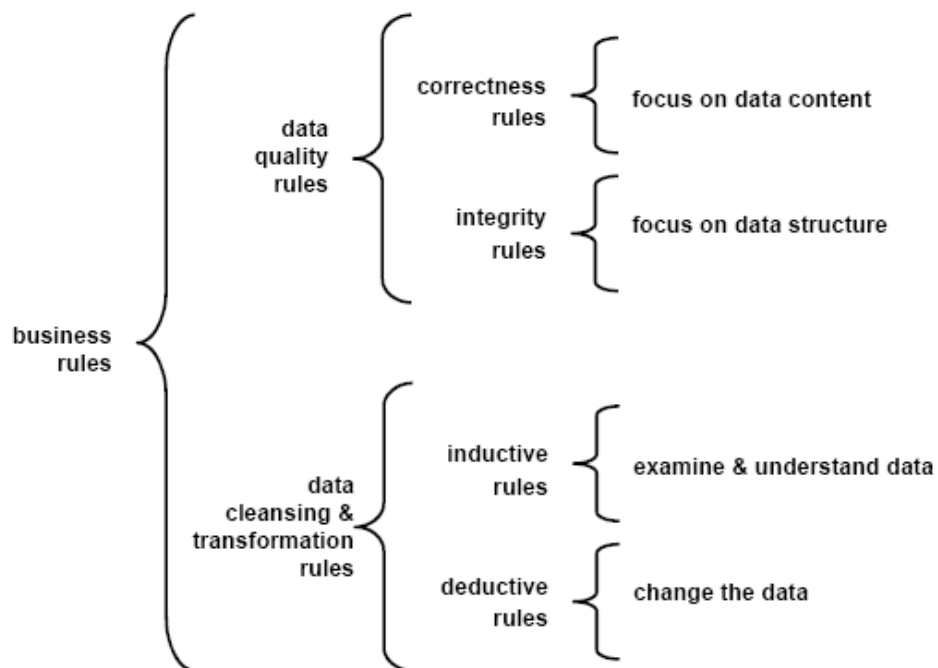
En segundo lugar, se deben especificar las reglas de limpieza y de transformación de datos. Las reglas de limpieza y de transformación de datos son reglas activas. Estas reglas no describen los defectos de los datos, sino las acciones que se tomarán cuando se descubren dichos defectos. Estas reglas se aplican para la reparación de defectos descubiertos en bases de datos y en la corrección de datos durante el gerenciamiento y la manipulación de los mismos.

Estas a su vez se subdividen en reglas inductivas y reglas deductivas.

Las reglas inductivas son aquellas que permiten examinar y entender los datos y las deductivas describen los cambios que se puedan llevar a cabo en estos.

La figura número 5 ilustra gráficamente el esquema anteriormente descrito.

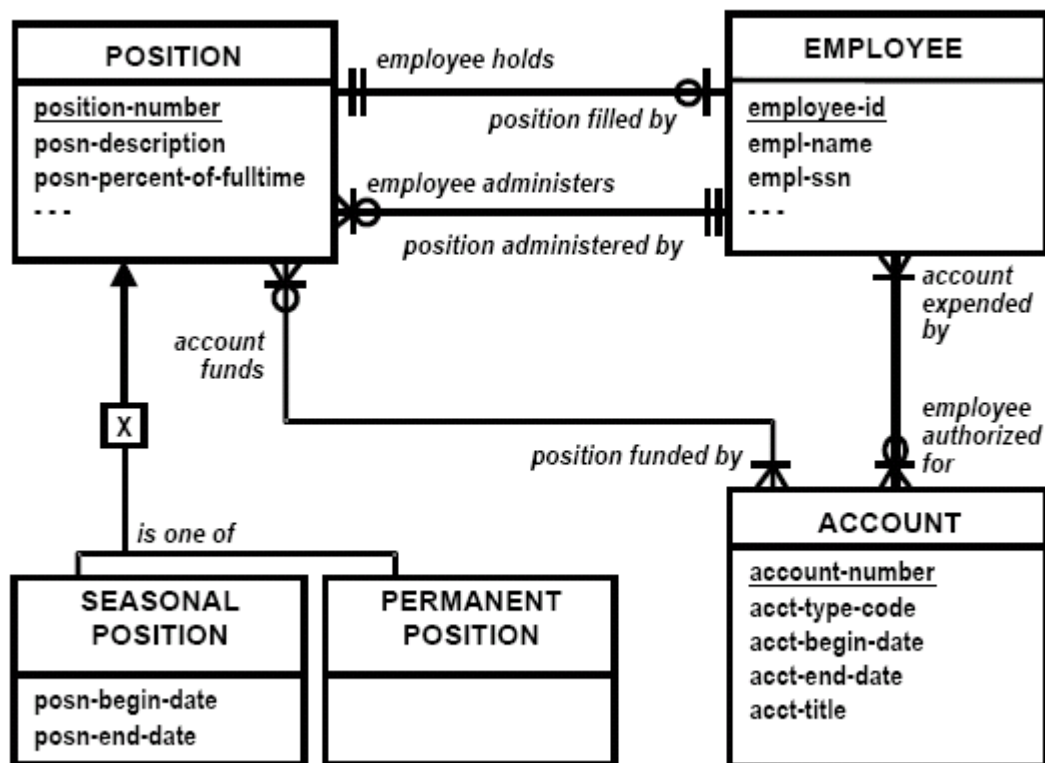
Figura 5. Calidad de Datos y Reglas del Negocio



Fuente: *DATA CLEANSING: Delivering high quality warehouse data*. J. Thomann

- **Fuente de datos:** en este campo se debe mencionar el canal que provee los datos en análisis. Para esta metodología aplica como canal el *Contact Center* y existen otros datos que son calculados por el sistema (Segmento, tipo de facturación, fecha de ingreso, etc.) o incluso existen datos que vienen de otros sistemas y se integran a la información del cliente.
- **Estructura de la base de datos:** Se especifica el modelo entidad relación de la base de datos y sus respectivas reglas de integridad, la relación de dependencia, descripción de los atributos, tipo de datos, cardinalidad, llaves primarias, etc.

Figura 6. Ejemplo de Modelo Entidad Relación



Fuente: *DATA CLEANSING: Delivering high quality warehouse data*. J. Thomann

- **Sistema CRM a impactar:** En este campo se deberá describir brevemente el aplicativo que se utiliza para mantener las relaciones con los clientes. La descripción deberá contener los siguientes aspectos:

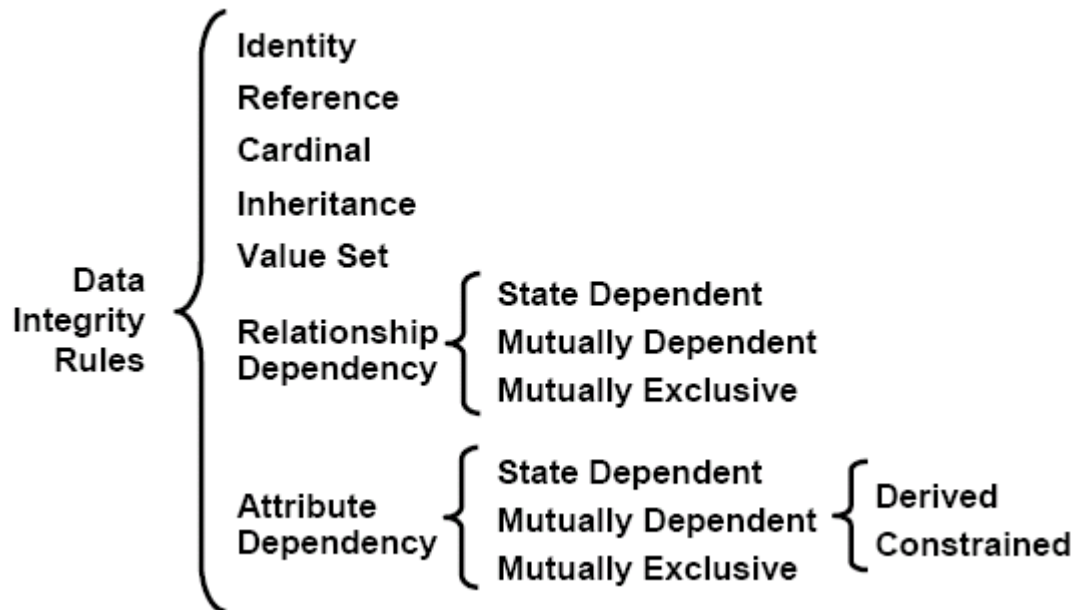
- Nombre del aplicativo.
- Descripción General de la funcionalidad.
- Áreas usuarios del aplicativo.
- Áreas impactadas
- Tipos de informes que se extraen de este.

- **Datos a evaluar:** De acuerdo a las reglas del negocio de la empresa cliente y a los atributos descritos en el modelo entidad relación, se deben listar los datos incluidos en el CRM a los cuales se les realizará el proceso de calidad de datos. Lo que a su vez debe determinar la prioridad que tiene cada campo para esta área, entre todos los campos que les interesan, algunos serán más importantes que otros.

La selección de estos datos se llevará a cabo de acuerdo a las necesidades de los procesos cliente del aplicativo, es decir, para el área de mercadeo será importante asegurar la calidad de datos en los campos “e-mail” para realizar una campaña a través de correo electrónico a nivel masivo, o la dirección de residencia para enviar una información publicitaria a un segmento de clientes determinados. Para el área de facturación será fundamental contar con la dirección correcta del cliente para el envío de su factura.

- **Reglas de integridad:** Las reglas de la integridad de datos, son esas reglas que gobiernan la estructura de los datos. La clasificación de las reglas de la integridad de datos se basa en el modelo Entidad-Relación y su clasificación se muestra en la figura 7 y se describen a continuación:

Figura 7. Reglas de Integridad de Datos



Fuente: *DATA CLEANSING: Delivering high quality warehouse data.* J. Thomann

- **Reglas de Entidad:** Estas garantizan que cada entidad tendrá una identidad única y asegura que los valores de la clave secundaria pueda referirse apropiadamente a valores de llave primaria. Toda tabla debe contar con una llave primaria y su valor es único. Ninguno de los atributos que componen la clave primaria puede ser nulo.

Como se observa en la figura 8 en la tabla *EMPLOYEE TABLE* se observa una violación a la regla de entidad ya que el campo *Employee ID number* tiene valores asociados a dos empleados diferentes, también se observa que un mismo empleado tiene asignados dos *ID Number* diferentes y campos nulos.

Figura. 8 Reglas de Entidad

EMPLOYEE TABLE

employee id number	name	status	hire date	separation date	position number
10271	Doe, John	ACTIVE	2001/03/16		9304
00391	Smith, Jan	RETIRED	1972/03/06	1997/03/05	3308
08794	Jones, Bob	SEPARATED	1996/10/17	2000/04/19	7714
• 11508	Amos, Jeff	ACTIVE	2000/01/02		6067
10299	Villareal,	ACTIVE	1984/07/12		7717
2 11622	Li, Yung	ACTIVE	2000/01/02		4056
10299	Smith, Joh	ACTIVE	1984/07/12		9017
• 11508	Orr, Wayne	ACTIVE	2000/01/02		6662
11262	Li, Yung	ACTIVE	2000/01/02		4056
3	Matthews,	ACTIVE	2000/01/02		6960

Fuente: *DATA CLEANSING: Delivering high quality warehouse data*. J. Thomann

- **Reglas de referencia:** Hace posible que un atributo no tenga un valor correspondiente, pero será imposible que tenga una entrada válida, la aplicación de la regla de integridad referencial hace posible eliminar una fila en una tabla cuya llave primaria contiene valores de la llave secundaria iguales en otras tablas.

En la figura 9 se ilustra una violación a esta regla, al no existir una referencia en la tabla *POSITION* de la posición número 9017 correspondiente a la tabla *EMPLOYEE*.

Figura.9 Reglas de Referencia

EMPLOYEE TABLE

employee id number	name	status	hire date	separation date	position number	
10271	Doe, John	ACTIVE	2001/03/16		9304	✓
00391	Smith, Jan	RETIRED	1972/03/06	1997/03/05	3308	✓
08794	Jones, Bob	SEPARATED	1996/10/17	2000/04/19	7714	✓
11508	Amos, Jeff	ACTIVE	2000/01/02		6067	✓
10299	Villareal,	ACTIVE	1984/07/12		7717	✓
11622	Li, Yung	ACTIVE	2000/01/02		4056	✓
10434	Smith, Joh	ACTIVE	1984/07/12		9017	?
11508	Orr, Wayne	ACTIVE	2000/01/02		6662	✓

POSITION TABLE

position number	position description	percent of fulltime	begin date	end date	type code
3308	Systems Mgr	100.00	1988/07/01		PERMANENT
9304	Gardener	060.00	2001/03/01	2001/09/15	SEASONAL
7717	Receptionis	100.00	2001/01/02		TEMPORARY
6067	Project Mgr	100.00	2001/01/02	2003/06/30	TEMPORARY
7714	Director of	100.00	2001/01/02		PERMANENT
4056	Database Ad	100.00	2001/01/02		PERMANENT
6662	Fiscal Tech	100.00	2001/01/02		PERMANENT
3138	Network Adm	100.00	2001/01/02		PERMANENT

Fuente: DATA CLEANSING: Delivering high quality warehouse data. J. Thomann

- **Reglas de Cardinalidad:** Describe la dimensión cuantitativa en la relación entre un par de entidades. La cardinalidad es una regla especial de integridad y refleja la política de la empresa.

En la figura 10 se muestran dos violaciones a la regla de cardinalidad, ya que en la tabla *EMPLOYEE*, al empleado 10271, no tiene una posición asignada. La regla

de negocio indica explícitamente que cada empleado debe tener asignada una posición. Igualmente se observa que los empleados 10299 y 10434 comparten una misma posición.

Figura.10 Reglas de Cardinalidad

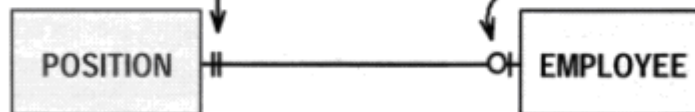
EMPLOYEE TABLE

employee id number	name	status	hire date	separation date	position number
10271	Doe, John	ACTIVE	2001/03/16		
00391	Smith, Jan	RETIRED	1972/03/06	1997/03/05	3308
08794	Jones, Bob	SEPARATED	1996/10/17	2000/04/19	7714
11508	Amos, Jeff	ACTIVE	2000/01/02		6067
10299	Villareal,	ACTIVE	1984/07/12		7717
11622	Li, Yung	ACTIVE	2000/01/02		4056
10434	Smith, Joh	ACTIVE	1984/07/12		7717
11508	Orr, Wayne	ACTIVE	2000/01/02		6662

Bussines Rules

*position filled by zero or one employees
employee fills exactly one position*

Data Model



POSITION TABLE

position number	position description	percent of fulltime	begin date	end date	type code
3308	Systems Mgr	100.00	1988/07/01		PERMANENT
9304	Gardener	060.00	2001/03/01	2001/09/15	SEASONAL
7717	Receptionist	100.00	2001/01/02		TEMPORARY
6067	Project Mgr	100.00	2001/01/02	2003/06/30	TEMPORARY
7714	Director of	100.00	2001/01/02		PERMANENT

Fuente: DATA CLEANSING: Delivering high quality warehouse data. J. Thomann

- **Valores permitidos:** esta regla de integridad hace referencia a las reglas del negocio que asignan parámetros validos para el contenido de los campos.

Por ejemplo: caracteres alfa numéricos, rangos de datos, uso de espacios, entre otras.

En la figura 11 se ilustran seis ejemplos de violación a esta regla:

1. La regla de negocio indica que el apellido debe ir separado por coma y un espacio del nombre.
2. El empleado 10299 tiene como fecha de cumpleaños 1902, valor que está por fuera del valor definido por la regla del negocio.
3. El empleado 10434 tiene como estado *ATIVE*, lo cual indica que la lista esta defectuosa.
4. En la tabla *POSITION*, la posición 9304 tiene un valor asignado que está por fuera del rango permitido por la regla del negocio.
5. La posición 6067 tiene un carácter no válido en la fecha.
6. La posición 4056 en la descripción tiene un carácter no permitido.

Figura 11. Valores Permitidos

EMPLOYEE TABLE

	employee id number	name	status	hire date	separation date	position number
1	10271	Poe John	ACTIVE	2001/03/16		9304
	00391	Smith, Jan	RETIRED	1972/03/06	1997/03/05	3308
	08794	Jones, Bob	SEPARATED	1996/10/17	2000/04/19	7714
	11508	Amos, Jeff	ACTIVE	2000/01/02		6067
2	10299	Villareal,	ACTIVE	1902/07/12		7717
	11622	Li, Yung	ACTIVE	2000/01/02		4056
3	10434	Smith, Joe	ACTIVE	1984/07/12		9017
	11508	Orr, Wayne	ACTIVE	2000/01/02		6662

POSITION TABLE

	position number	position description	percent of fulltime	begin date	end date	type code
	3308	Systems Mgr	100.00	1988/07/01		PERMANENT
4	9304	Gardener	600.00	2001/03/01	2001/09/15	SEASONAL
	7717	Receptionist	100.00	2001/01/02		TEMPORARY
5	6067	Project Mgr	100.00	2001/01/02	2001/06/30	TEMPORARY
	7714	Director of	100.00	2001/01/02		PERMANENT
6	4056	Database Ad	100.00	2001/01/02		PERMANENT
	6662	Fiscal Tech	100.00	2001/01/02		PERMANENT
	3138	Network Adm	100.00	2001/01/02		PERMANENT

Business Rules

- list of allowable values
- range of allowable values
- allowable character set
- pattern mask
- combination of above

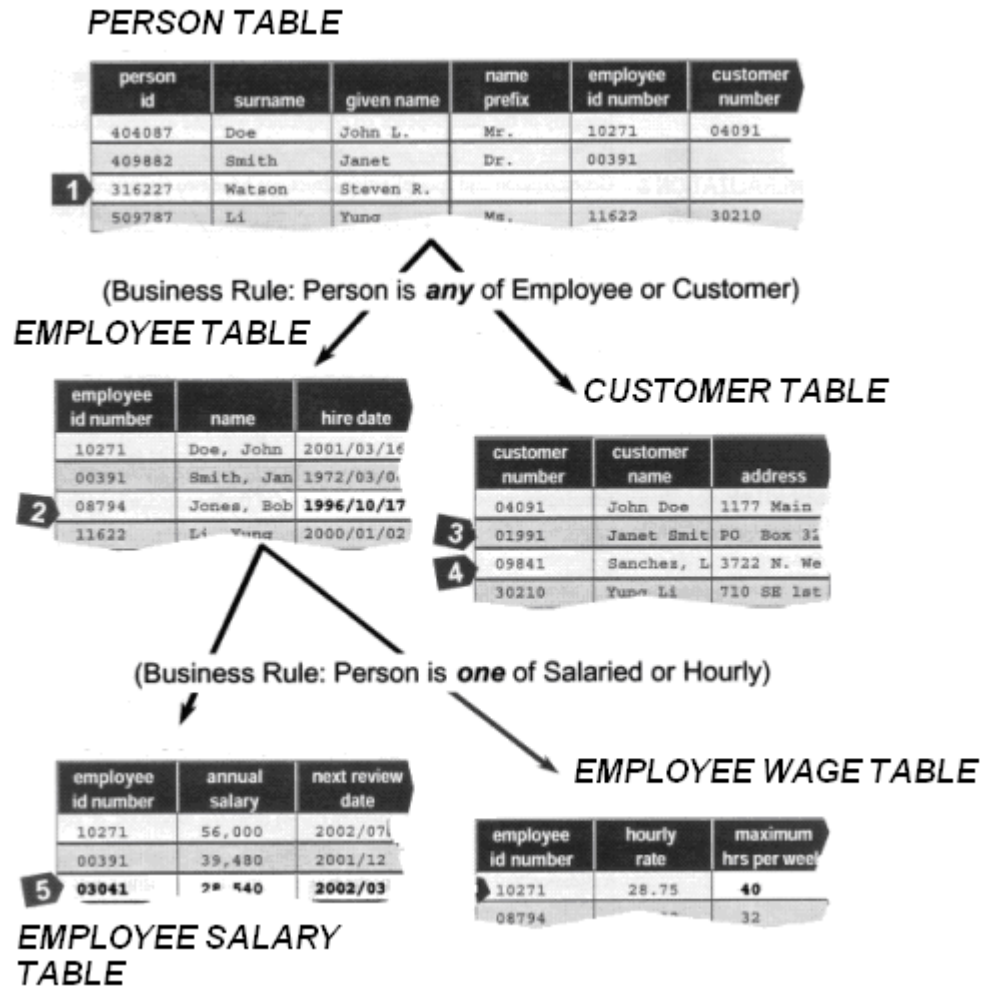
Fuente: DATA CLEANSING: Delivering high quality warehouse data. J. Thomann

- **Reglas de Herencia:** Es una propiedad que permite que los objetos sean creados a partir de otros ya existentes, obteniendo características (métodos y atributos) similares a los ya existentes. La herencia es un tipo de relación entre una entidad "padre" y una entidad "hijo". La entidad "hijo" hereda todos los atributos y relaciones de la entidad "padre".

La figura 12 muestra siete violaciones a las reglas de herencia, las cuales son:

1. La persona identificada con el ID 316227 no es ni empleado ni cliente, lo cual viola a regla de negocio que indica que toda persona debe ser o empleado o cliente.
2. El empleado 08794 existe en la tabla *EMPLOYEE* pero no en la tabla persona. La regla de negocio indica que un subtipo debe corresponder al supertipo, en este caso la tabla persona.
3. El cliente 01991 no aparece relacionado en la tabla persona, pero si en la tabla empleado con el ID 00391, esto es un defecto del mapeo de la llave. Esta verificación se lleva a cabo para asegurar que una misma persona es empleado y cliente.
4. El cliente 09841 existe en la tabla cliente, pero no en la tabla persona, es decir, el subtipo no aparece en el supertipo.
5. El empleado 03041 existe como subtipo en la tabla *SALARY EMPLOYEE*, pero no existe como supertipo en la tabla *EMPLOYEE*.

Figura 12. Reglas de Herencia



Fuente: DATA CLEANSING: Delivering high quality warehouse data. J. Thomann

- **Relación de Dependencia:** Ocurre cuando la existencia de una relación es condicional y las condiciones bajo las cuales existe dependen del estado de una entidad que participa con la existencia de otra relación.

Figura 13. Relación de Dependencia

EMPLOYEE TABLE

employee id number	name	status	hire date	separation date	position number
10271	Doe, John	ACTIVE	2001/03/16		9304
00391	Smith, Jan	RETIRED	1972/03/06	1997/03/05	3308
08794	Jones, Bob	SEPARATED	1996/10/17	2000/04/19	7714
11508	Amos, Jeff	ACTIVE	2000/01/02		6067
10299	Villareal,	ACTIVE	1984/07/12		7717
11622	Li, Yung	ACTIVE	2000/01/02		4056
10434	Smith, Joh	ACTIVE	1984/07/12		9017
11508	Orr, Wayne	ACTIVE	2000/01/02		6662

(Business Rule: Only active employees may participate in the investment program)

Investment Program Table

	employee id number	deduction amount	deduction period
	11622	425.00	Monthly
1	08794	185.00	Monthly
	10271	1000.00	Quarterly
2	00391	184.00	Monthly

Fuente: DATA CLEANSING: Delivering high quality warehouse data. J. Thomann

- **Atributos de Dependencia:** Las reglas de la dependencia de los atributos se presentan cuando el sistema de los valores permitidos para un atributo está parcialmente determinado por los valores de uno o más atributos. Existen tres tipos de dependencia de atributos:

La dependencia de estado: Se presenta cuando los valores de los atributos dependen del estado que tiene la entidad y del valor de los atributos que describen ese estado.

Dependencia mutua: Existe cuando el valor de los atributos por lo menos determina parcialmente los valores de otros atributos. La dependencia mutua entre los atributos se presenta cuando:

Hay derivación, es decir, cuando el valor correcto de un atributo puede determinar el valor de otro atributo.

Y cuando el valor permisible para un atributo se puede determinar parcialmente de los valores de otros atributos; o cuando el sistema de valores permitidos se puede reducir a un subconjunto basado en los valores de otros atributos.

Exclusividad Mutua: existe cuando el valor de un atributo depende de otros atributos de valor nulo.

En la figura 14 se ilustran cuatro defectos de la dependencia de atributos:

1. El empleado 00391 tiene una fecha de separación y la fecha próxima de la revisión del sueldo. Éstos son mutuamente - datos exclusivos.
2. El empleado 08794 tiene una fecha de la separación que preceda su fecha del alquiler. Esto viola una restricción de dependencia mutua.
3. El empleado 10299 esta ACTIVO pero tiene una fecha de separación. Esto viola una restricción de la dependencia de estado.
4. Un registro de la historia salarial para el empleado 10299 tiene FTE = 1.000, sueldo de tiempo completo anual = 56.220 y tarifa de paga anual = 42.165 (el 75% del sueldo a tiempo completo). Esto viola una restricción de dependencia entre datos derivables.

Figura 14. Atributos de Dependencia.

EMPLOYEE TABLE

	employee id number	name	status	hire date	separation date	next salary review date
1	10271	Doe, John	ACTIVE	2001/03/16		2002/01/01
2	00391	Smith, Jan	RETIRED	1972/03/06	1997/03/05	2002/01/01
3	08794	Jones, Bob	SEPARATED	1996/10/17	1994/04/19	
	11508	Amos, Jeff	ACTIVE	2000/01/02		2001/07/01
	10299	Villareal, Li	ACTIVE	1984/07/12	2001/01/05	2002/01/01
	11622	Li, Yung	ACTIVE	2000/01/02		2002/07/01
	10434	Smith, John	ACTIVE	1984/07/12		2002/01/01
	11508	Orr, Wayne	ACTIVE	2000/01/02		2001/10/01

FISCAL YEAR SALARY HISTORY TABLE

	employee id number	salary date	FTE	annual full- time salary	annualized pay rate	salary grade
4	10299	2000/06/30	1.000	57,780	57,780	11B
	10299	1999/06/30	1.000	56,220	42,165	11B
	10299	1998/06/30	1.000	46,224	46,224	10C
	10299	1997/06/30	1.000	38,304	38,304	8C
	10299	1996/06/30	0.800	36,000	28,800	8A
	10299	1995/06/30	0.800	31,500	25,200	6B
	10299	1994/06/30	0.800	28,462	22,770	5B
	10299	1993/06/30	0.500	27,000	13,500	5B
	10299	1992/06/30	0.500	24,360	12,180	4C
	11622	2000/06/30	1.000	29,000	29,040	5B
	11622	1999/06/30	1.000	26,000	26,400	5A

Fuente: DATA CLEANSING: *Delivering high quality warehouse data*. J. Thomann

3.1.2 Objetivos e indicadores: El proceso de la planeación debe establecer objetivos e indicadores de medición de la corrección de datos, estos indicadores permitirán medir los impactos alcanzados con la ejecución del proyecto de la calidad de datos.

Objetivos pueden estar descritos en dos formas diferentes:

- Expectativa relacionada con el logro en términos de la percepción y satisfacción de los usuarios de la base de datos.
- Logros en la eliminación de errores.

Para establecer los objetivos y definir los indicadores que se utilizarán para medir se debe tener claro los datos seleccionados anteriormente para aplicarles la metodología de calidad de datos y comenzar a evaluar la base de datos de las empresas clientes y verificar el estado de los datos en un tiempo determinado y así realizar el seguimiento adecuado teniendo en cuenta los siguientes indicadores:

- **Indicador 1:** *Porcentaje de Variación de la base de datos de clientes*, este indicador se desarrolla extrayendo del sistema CRM el comportamiento de la base de datos en un tiempo determinado, en aumento o en disminución de registros de información de los clientes. El objetivo es conocer el total de la base de datos que se maneja en la empresa cliente y el comportamiento en un período de tiempo determinado, conociendo así el impacto de los diferentes eventos sobre la base total de clientes.

Este indicador se calcula con la siguiente fórmula:

$$\% \text{ de variación base de datos} = \# \text{ Registros actuales} / \# \text{ registros anteriores}$$

- **Indicador 2:** *Porcentaje de Campos diligenciados correctamente*: este indicador mide sobre el total de registros de clientes en la base de datos, el total de clientes que quedaron con determinado dato diligenciado correctamente y por ende cuantos quedaron con el campo nulo o mal diligenciado. El objetivo de este indicador determinar la cantidad de campos faltantes o errados que están en la base de datos, detectando los errores más frecuentes y corrigiéndolos reactivamente a corto plazo.

Es importante tener en cuenta que cuando se habla de “Campos mal diligenciados” es una aproximación, pues lo que se hace es que se programa el reporte para que muestre los errores ya conocidos y más comunes, pero los nuevos errores que se generen en la base de datos no serán medibles en este informe y se quedarán por fuera de este indicador.

% campos correctos = # de registros correctos / # registros totales BD

- **Indicador 3:** *Registros corregidos*, se medirán constantemente los campos corregidos sobre los campos errados, hasta llegar a un 100% de corrección, previniendo así que queden datos errados en el sistema y el crecimiento del error en el tiempo que dure el desarrollo de la corrección del error. Su objetivo es medir la efectividad de las estrategias desarrolladas para la corrección de los campos detectados con errores.

% Registros corregidos = # registros corregidos / # total de registros errados

El anexo B. “Objetivos e Indicadores de la Calidad de Datos” se debe consignar la información de los resultados de la medición de los tres indicadores anteriormente mencionados mes a mes para realizar el seguimiento de la calidad de datos la cual apoyará a la toma de decisiones del proyecto.

Igualmente se cuanta con un espacio para detallar los hallazgos detectados a raíz de los resultados y análisis de los mismos y los planes de acción a emprender.

3.1.3 Medidas de Calidad: Cuantifica la calidad de los datos con el fin de crear confianza entre los usuarios de la base de datos.

La medición es necesaria para evaluar objetivamente el estado actual de la calidad de los datos y el progreso de los objetivos planeados.

La medición de los indicadores facilita contar con datos más exactos, lo que genera menos subjetividad para la evaluación del estado actual de los datos, permitiendo que el análisis sea más objetivo entre la realidad y los objetivos propuestos.

Analice cada uno de los resultados obtenidos con la medición de los indicadores definidos para la calidad de datos y evalúe el estado actual respecto a la meta establecida. Con el resultado de este análisis se deberá definir el tipo de acciones que se deberán tomar para disminuir la brecha entre lo real y lo esperado en el indicador que se este analizando.

En el anexo C “Medidas de Calidad” se ingresa la información de las metas y del comportamiento del indicador y se calcula la brecha existente entre estas dos variables. A partir de este cálculo se realiza el análisis para la toma de acciones.

3.1.4 Acciones: Se debe describir las diferentes acciones de mejora que se realizarán para cumplir con el objetivo.

Es importante para elegir el tipo de acción a implementar, tener muy claras las respuestas a los siguientes interrogantes:

¿Cuál es el plan para detectar errores?

¿Con qué plan se corrige cuando los encuentra?

¿Cuál es el plan para prevenir errores en las fuentes?

Existe una relación entre las reglas de integridad y las acciones a tomar, ya que a partir de las reglas de integridad que posee la fuente de datos se especifican las acciones a realizar cuando se detecta una violación a cualquiera de dichas reglas.

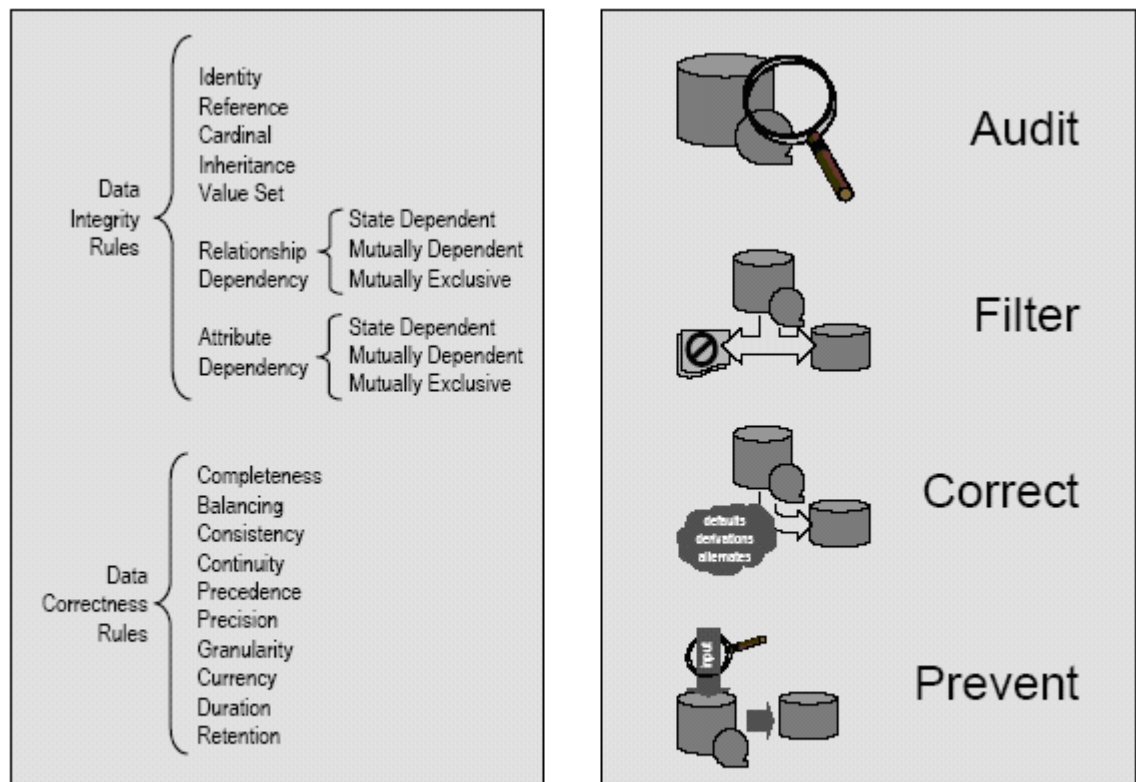
Los procedimientos de limpieza son los primeros pasos para alcanzar la calidad de los datos y estos deben basarse en las reglas del negocio.

Existen diferentes tipos de acciones, las principales son:

- Auditoría, filtros y correcciones a la fuente de datos.
- Auditoría, filtros y correcciones a los campos en blanco.
- Prevención de los errores en la fuente.

La figura 15. Ilustra la relación entre las reglas de integridad y de corrección y las diferentes tipos de acciones a tomar.

Figura 15. Definición de Acciones



RULES + ACTIONS

Fuente: *DATA CLEANSING: Delivering high quality warehouse data*. J. Thomann

A continuación se describe cada una de las acciones anteriormente mencionadas:

- **Auditoría:** El propósito de la revisión es conocer el grado de los problemas existentes en la calidad de los datos, el grado y su severidad. La auditoría es el primer paso esencial de la calidad de los datos, esta examina el contenido, la estructura, el grado de completación, la distribución de valores, entre otros factores que detectan y permiten conocer si existe una violación a las reglas de la integridad y de la corrección. El entregable de la auditoría será un informe descriptivo y estadístico con los hallazgos encontrados y los aspectos a mejorar.
- **Filtros:** El propósito de los filtros es quitar datos problemáticos de un sistema de datos. Esta acción es típicamente aplicada cuando los datos se consideran defectuosos a un grado tal que los haga inutilizables. Existen diferentes tipo de filtros:

Definir datos nulos: Se debe analizar los campos que no contengan información (nulos), dependiendo de esta análisis de les asigna un valor predeterminado o se eliminan.

Eliminación de registros duplicados: Se deben identificar los registros duplicados con el fin de garantizar las reglas de no duplicación.

Análisis lógico de campos: El grupo de campos se deben analizar de acuerdo al tipo de dato alojado en él. Por ejemplo: en el campo fecha deben existir realmente fechas, no números de cédula.

Estos filtros se llevan a cabo en la fuente de datos y tienen como objetivo evitar que sean almacenados datos inconsistentes

- **Correcciones:** Se deben detectar los datos defectuosos, con el fin de sustituirlos por el dato correcto de acuerdo al tipo de campo. Existen varias técnicas para corregir datos:

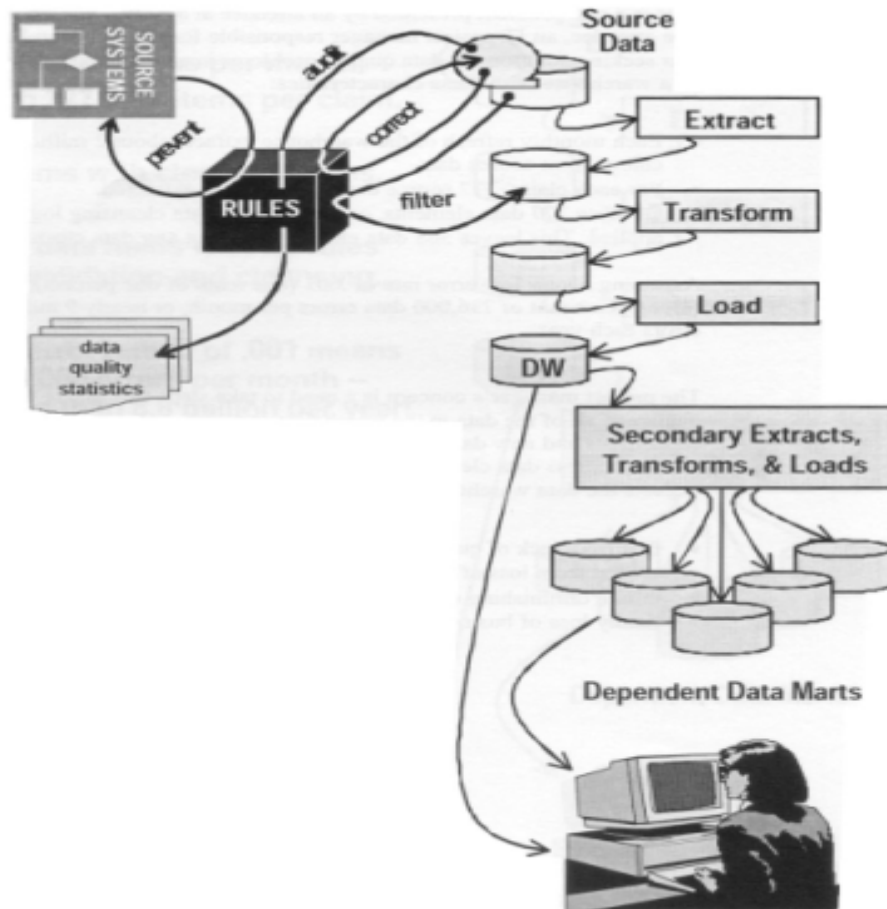
Insertando un valor predefinido que indica la ausencia de datos confiables. Por ejemplo: cuando el campo fecha de nacimiento no es confiable, se puede reemplazar el valor por una etiqueta que diga “Por Confirmar”.

Detección de errores comunes de la entrada de datos y se determinan medidas para minimizarlos. Por ejemplo: valores predeterminados de datos en listas, es decir, en el campo género solo se podrá elegir F para femenino y M para masculino.

Búsqueda de fuentes alternativas para encontrar un valor de reemplazo. Por ejemplo: el cruce con otras bases de datos.

- **Prevención:** el propósito de la prevención es eliminar las causas de la aparición de datos defectuosos y la toma de medidas para evitar que se vuelvan a presentar.

Figura 16. Ciclo de Vida de la Calidad de Datos



Fuente: *DATA CLEANSING: Delivering high quality warehouse data.* J. Thomann

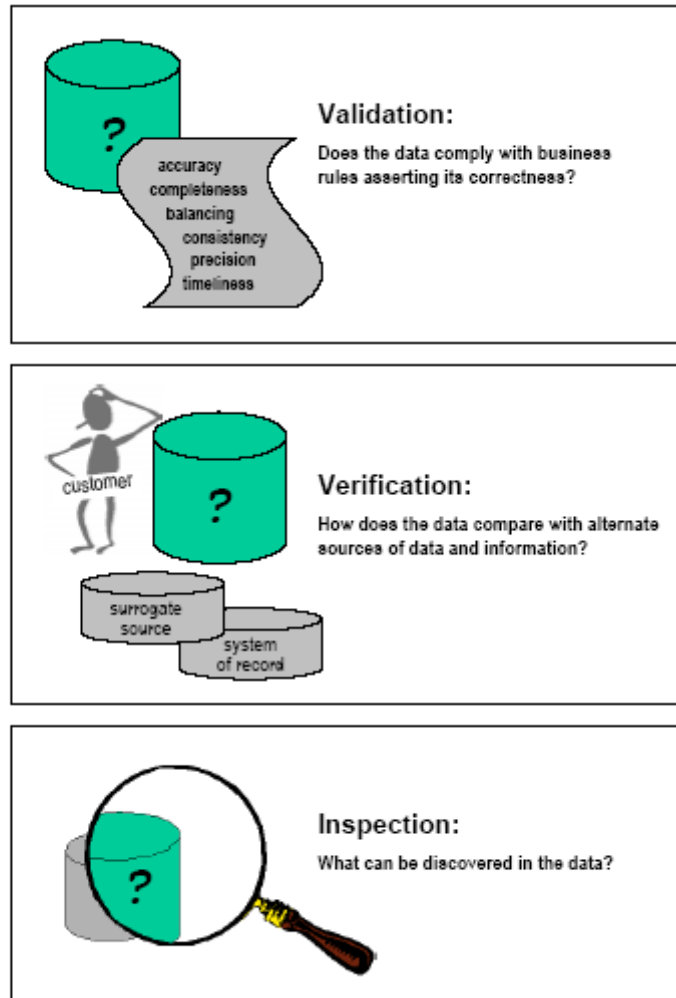
La figura 16. Muestra el ciclo de vida de la calidad de datos. El cual comienza identificando las reglas de la calidad de los datos y empaquetándolas como componentes, luego se debe auditar la fuente de datos y comenzar a monitorear la calidad de los datos. Posteriormente, después se extrae la información se filtra y/o corrige, lo cual permite identificar oportunamente los errores. Finalmente, prevenga los errores en la fuente de datos utilizando las reglas de corrección.

Después de desarrollar el plan de acciones para la corrección de datos, se debe realizar un proceso de chequeo de los resultados obtenidos.

Este proceso contempla tres fases:

- **Validación:** prueba de validación contra las reglas de negocio definidas para el dato en análisis.
- **Verificación:** Prueba de verificación contra fuentes fiables, estas fuentes fiables pueden ser del “mundo real”, como un cliente específico o un cruce contra otra base de datos ya existente. La primera es mucho más costosa, pero le agregue un porcentaje muy alto de confiabilidad y eficiencia a la verificación y la segunda es más económica y se puede llevar a cabo automáticamente.
- **Inspección:** La inspección conduce una examinación cuidadosa de datos para descubrir las características que no se pudieron encontrar usando técnicas de la validación y de verificación. Donde la validación y la verificación se basan en criterios de inspección de reglas de negocio y de las fuentes alternativas.

Figura17. Chequeando la corrección de datos.



Fuente: *DATA CLEANSING: Delivering high quality warehouse data*. J. Thomann

Posteriormente es necesario realizar un proceso de detección de defectos de la corrección de los datos, para ello es necesario tener en cuenta las siguientes técnicas básicas:

- **Exactitud:** Los datos son exactos cuando coinciden con una fuente confiable de información. La prueba de la exactitud evalúa los datos contra uno o varias fuentes confiables. La fuente de chequeo utilizada en la verificación.

- **Completación:** Todos los datos deben estar diligenciados. La prueba de ello son filas incompletas y datos con valores nulos cuando se espera un valor. Las fuentes de chequeo utilizadas son la verificación, la validación e inspección.
- **Balance:** Los datos numéricos deben conservar unos márgenes de error permitidos, por lo tanto cuando se presenten variaciones, estas deberán estar dentro del rango permitido. La fuente de chequeo utilizada es la verificación.
- **Consistencia:** Los datos son consistentes cuando cumplen con las características definidas para el campo específico. Las fuentes de chequeo utilizadas son la verificación, la validación.
- **Precisión:** Los datos son exactos cuando se registran con el grado de exactitud que se requiere de acuerdo a las necesidades y expectativas de información del negocio. Precisión la prueba evalúa el nivel de exactitud para detectar los datos no suficientemente exactos para sus aplicaciones previstas. Las fuentes de chequeo utilizadas son la verificación, la validación e inspección.
- **Ocurrencia:** La cantidad de diferencia de tiempo entre los datos actuales y la fecha pasada en que los datos fueron cargados en la base de datos. La prueba determina la cantidad de tiempo que ha transcurrido desde que los datos fueron cargados por última vez para cualquier asunto del interés. Las fuentes de chequeo utilizadas son la validación e inspección.
- **Duración:** Es el rango de tiempo para el cual un grupo de datos representa el mundo real. La duración se aplica a los datos con una fecha inicio y una fecha fin. Prueba de la duración identifica duraciones ilógicas o ambiguas. Las fuentes de chequeo utilizadas son la validación e inspección.

- **Retención:** Es el período de tiempo que los datos están guardados en la base de datos. Los datos pueden ser retirados por vencimiento o almacenamiento fuera de línea. La prueba de la retención identifica los datos más viejos de una tabla. La fuente de chequeo utilizada es la validación
- **Precedencia:** Los datos son precedentes cuando presentan una secuencia histórica que se ajusta a las reglas de negocio. La prueba de la precedencia examina datos para detectar secuencias de evento prohibidas o ilógicas. Las fuentes de chequeo utilizadas son la verificación, la validación.

3.1.5 Recursos y Responsabilidades: Definir los recursos y responsabilidades necesarias para detectar, corregir y prevenir defectos en la calidad de los datos.

Se debe realizar un plan completo, donde se identifiquen los recursos y responsabilidades para medir y monitorear la efectividad del plan. En esta metodología se proponen cuatro equipos de trabajo los cuales desempeñarán diferentes roles para la calidad de datos:

- **Líder:** es la persona que entiende el impacto de la calidad de datos, desarrolla estrategias, políticas de calidad, se encarga de liderar los cambios pertinentes que se deben realizar en los datos, presenta informes a las empresas clientes para realizar seguimiento al proceso.

Como el “Oficial Principal de Calidad” (CDQ). Monitorea los niveles de calidad, analiza las fuentes de problemas y recomienda acciones correctivas.

- **Dueño del Proceso:** son las personas responsables de todo el ciclo básico de conocimiento, y adicionalmente administran todo el proceso.

Son todos aquellos responsables directos de los procesos que crean, modifican y transportan datos.

- **Profesionales de Información:** son las personas con la responsabilidad de administrar la información, “Oficiales principales de Información” (CDI). Son quienes trabajan con los datos, los convierten en información y presentan los resultados, realizan transferencia de conocimiento y soportan los cambios.
- **Personal Operativo *Contact Center*:** son las personas encargadas de capturar la información de los clientes e ingresarla en el CRM. La calidad de datos parte de esta interacción.

3.1.6 Cronograma: Se debe definir un plan de trabajo que contenga las prioridades para desarrollar el proyecto de calidad de datos. Este cronograma incluye fechas de entrega y un esquema de seguimiento que permita medir el avance del proyecto.

Es importante recordar que el cronograma hace referencia al proyecto de montaje de la metodología, pero el proceso de calidad de datos debe ser continuo por lo tanto no se debe manejar un cronograma específico.

El anexo D. es un formato de cronograma que puede ser utilizado como guía para el desarrollo de la metodología de calidad de datos.

3.1.7 Continuidad: Un plan de Calidad de Datos busca ser permanente. Continuamente la calidad está encontrando nuevas mejoras y nuevos cambios que deben implementarse y monitorearse.

La continuidad para esta metodología no es simplemente ejecutar acciones de limpieza de datos, no es suficiente asegurarse de que la calidad de los datos

mejora a medida que se evalúa regularmente, es de suma importancia cuantificar el impacto de las acciones de limpieza y asignar el responsable de monitorear los indicadores y supervisar los efectos de las acciones de limpieza de los datos.

Se debe tener las siguientes premisas para darle continuidad al proceso.

- Medir por medio de indicadores.
- Adaptar y ajustar.
- Valorar de nuevo el estado actual
- Determinar nuevas metas.
- Definir nuevo planes de acción.

4. RECOMENDACIONES

Como proceso de apoyo a todo el desarrollo de calidad de datos se debe contar con un espacio en donde se compartan todos los indicadores con las áreas impactadas, pues el proceso de calidad de datos y sus mediciones no se ejecutan sólo para que el responsable de la calidad de datos esté enterado, tanto las áreas de sistemas, las áreas operativas del *Contact Center*, mercadeo y el área encargada de procesar la información deben conocerlos para identificar sus aportes dentro del proceso; además esta práctica ayuda mucho a concienciar a la empresa en la importancia de la calidad de datos.

Asimismo, se debe contar con un proceso debidamente documentado en donde se especifique la forma en que el responsable de calidad de datos le va a solicitar al área técnica un ajuste sobre la base de datos, por lo general estos son unos formatos que se diligencian bajo la supervisión de muchas áreas, pues tocar directamente las bases de datos de los clientes es un tema de mucho riesgo y debe estar muy claro el propósito y quien autoriza el cambio.

Como esta metodología está enfocada en el *Contact Center*, es fundamental diseñar un esquema de divulgación muy bien definido, para garantizar pese a la alta rotación del personal, la continuidad del proyecto de la calidad de datos. Este esquema deberá apoyarse por el área de formación de Emtelco S.A.

BIBLIOGRAFÍA

[Marketing Directo, 2004] La Baja Calidad De Los Datos: Un Mal Común
<http://www.marketingdirecto.com/noticias/noticia.php?idnoticia=11002>

[Marketing Directo, 2006] Schober PDM y ESIC Business & Marketing School,
Noticias Bases de Datos y CRM, Artículo El 74% De Las Empresas Falla En
Calidad De Datos
<http://www.marketingdirecto.com/noticias/noticia.php?idnoticia=20870>

[REDMAN, 1996] Redman, Thomas C., *"Data for the Information Age"*, 1996.
Arthec House, Inc. Norwood, MA 02062.

DATA CLEANSING: Delivering high quality warehouse data. J. Thomann. *The Data Warehouse Institute.* 2002

Presentación Calidad de Datos Orbitel S,A E.S.P. Dirección de Gestión de la Información. Andrés Orozco O. Mariluz Tejada Z. 2006.

Presentación Planteamiento estratégico Emtelco S.A. Dirección General. 2007.

Publicado por Jesús Hoyos el 15/02/07 a las 6:15 en Calidad de Datos, Lealtad
http://www.jesushoyos.typepad.com/crm_en_latinoamerica/2007/02/calidad_de_datos.html

ANEXO A. ENUNCIADO PRELIMINAR DEL ALCANCE DE LA CALIDAD DE DATOS

Nombre de la Empresa:
Elaborado por:
Fecha:

Descripción general del negocio:	
Reglas del negocio:	1. Reglas de Calidad de Datos: 1.1. Reglas de corrección: 1.2 Reglas de integridad:
	2. Reglas de limpieza y transformación de datos: 2.1 Reglas inductivas: 2.2 Reglas deductivas
Fuente de datos:	Contact Center

Estructura de la base de datos:	Modelo Entidad Relación (gráficamente). Se deben incluir el tipo de dato, cardinalidad, llaves primarias.
Sistema impactar: CRM a	<ul style="list-style-type: none"> - Nombre del aplicativo: - Descripción General de la funcionalidad: - Áreas usuarios del aplicativo: - Áreas Impactadas: - Tipos de informes que se extraen de este:
Datos a evaluar:	

ANEXO B. OBJETIVOS E INDICADORES DE LA CALIDAD DE DATOS

Nombre de la Empresa:
Elaborado por:
Fecha:

INDICADOR / MES	ENERO	FEBRERO	MARZO	ABRIL	MAYO	JUNIO	JULIO	AGOSTO	SEPTIEMBRE	OCTUBRE	NOVIEMBRE	DICIEMBRE
Indicador 1: <i>% Variación de la base de datos de clientes</i>												
Indicador 2: <i>% de Campos diligenciados correctamente</i>												
Indicador 3: <i>Registros corregidos</i>												
Indicador 4:												
Indicador 5:												

Meses	Análisis de indicadores
Enero	Hallazgos
	Planes de Acción
	Áreas Impactadas
Febrero	Hallazgos
	Planes de Acción
	Áreas Impactadas
Marzo	Hallazgos
	Planes de Acción
	Áreas Impactadas
Abril	Hallazgos
	Planes de Acción
	Áreas Impactadas
Mayo	Hallazgos
	Planes de Acción
	Áreas Impactadas
Junio	Hallazgos
	Planes de Acción
	Áreas Impactadas
Julio	Hallazgos
	Planes de Acción
	Áreas Impactadas
Agosto	Hallazgos
	Planes de Acción
	Áreas Impactadas
Septiembre	Hallazgos
	Planes de Acción
	Áreas Impactadas
Octubre	Hallazgos
	Planes de Acción
	Áreas Impactadas
Noviembre	Hallazgos
	Planes de Acción
	Áreas Impactadas
Diciembre	Hallazgos
	Planes de Acción
	Áreas Impactadas

ANEXO C. MEDIDAS DE CALIDAD

Nombre de la Empresa:
Elaborado por:
Fecha:

Fecha	Indicador	% Variación de la base de datos de clientes	% de Campos diligenciados correctamente	% Registros corregidos	Acciones
Enero	Meta				
	Resultado				
	Brecha				
Febrero	Meta				
	Resultado				
	Brecha				
Marzo	Meta				
	Resultado				
	Brecha				
Abril	Meta				
	Resultado				
	Brecha				
Mayo	Meta				
	Resultado				
	Brecha				
Junio	Meta				
	Resultado				
	Brecha				

Fecha	Indicador	% Variación de la base de datos de clientes	% de Campos diligenciados correctamente	% Registros corregidos	Acciones
Julio	Meta				
	Resultado				
	Brecha				
Agosto	Meta				
	Resultado				
	Brecha				
Septiembre	Meta				
	Resultado				
	Brecha				
Octubre	Meta				
	Resultado				
	Brecha				
Noviembre	Meta				
	Resultado				
	Brecha				
Diciembre	Meta				
	Resultado				
	Brecha				

ANEXO D.
CRONOGRAMA NOMBRE PROYECTO CLIENTE

Versión No.
Fecha Actualización:

Actividades	Duración	Comienzo	Fin	% Completado	Recursos
Documento Diagnostico situación actual					
Toma de muestra					
Realizar listado de clientes activos en el <i>contact center</i>					
Determinar los criterios de selección de la muestra					
Seleccionar muestra.					
Toma de muestra finalizada					
Evaluación muestra					
Determinar criterios de evaluación de la muestra.					
Evaluar muestra					
Determinar resultados de la evaluación					
Evaluación muestra finalizada					
Selección de Datos					
Determinar criterios de selección de datos					
Seleccionar los datos					
Selección de datos finalizado					
Resultado de análisis					
Determinar mejoras en los datos evaluados					
Resultado del análisis finalizado					
Documento de Diagnostico situación actual finalizado					
Documento planeación					
Definición del alcance					
Determinación objetivos e indicadores					
Definición medidas de Calidad					
Definición acciones					

Actividades	Duración	Comienzo	Fin	% Completado	Recursos
Selección de recursos y responsabilidades					
Definición cronograma					
Definición de la continuidad					
Documento planeación finalizado					
Documento preliminar de la metodología					
Estructura general de la metodología					
Conocer las reglas del negocio					
Identificar las fuentes de los datos					
Conocer la estructura de la base de datos					
Seleccionar los datos a evaluar					
Determinar los datos y reglas que serán excluidos					
Establecer objetivos e indicadores de medición de la corrección de datos					
Evaluar objetivamente el estado actual de la calidad de los datos y el progreso de los objetivos planeados					
Generar la estrategia para detectar errores					
Detallar el plan de corrección de errores detectados					
Definir la estrategia de prevención de errores en las fuentes					
Definir los recursos y responsabilidades necesarias para detectar, corregir y prevenir defectos en la calidad de los datos					
Estructura general de la metodología finalizada.					
Cronograma metodología					
Identificar la duración de cada una de las etapas de la metodología					
Definir los responsables de cada etapa					
Definir cronograma con las prioridades para desarrollar el proyecto de calidad de datos.					
Cronograma metodología finalizado					
Documento preliminar de la metodología terminado					
Documento Revisado					

Actividades	Duración	Comienzo	Fin	% Completado	Recursos
Estructura general revisada					
Validar cada uno de los componente de la metodología					
Sugerir cambios o correcciones					
Estructura general revisada terminada					
Cronograma revisado					
Validar cada componente del programa (tiempos, responsables y actividades)					
Sugerir cambios o correcciones					
Cronograma revisado terminado					
Documento Revisado terminado					
Documento final					
Llevar a cabo los cambios y correcciones solicitados.					
Documento final terminado					
Documento de Cierre					
Elaborar y firmar el acta de entrega					
Documento de Cierre terminado					
Documento de Gestión					
Recolección de documentación					
Recopilar información reuniones, informes de avances, requerimientos de cambios, entre otras.					
Recolección de documentación finalizada					
Documento de Gestión terminado					